

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:02 ; Search time 118.14 Seconds  
(without alignments)  
235.123 Million cell updates/sec

Title: US-09-596-746A-24

Perfect score: 1842

Sequence: 1 MACNNSGKDGNTSANSADSE.....KNLSKAKEMLTNSVKELTS 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq-1101.\*

1: /SID8/gcgdata/geneseq/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1837	99.7	374	22	AAB62710
2	1834	99.6	398	22	AAB62728
3	1819.5	98.8	401	22	AAB62738
4	1819.5	98.3	378	22	AAB62712
5	1592.5	86.5	373	22	AAB62711
6	1589.5	86.3	397	22	AAB62729
7	1575	85.3	400	22	AAB62739
8	1566	85.0	377	22	AAB62713
9	1529	83.0	370	22	AAB62715
10	1526	82.8	394	22	AAB62730
11	1493	81.1	368	22	AAB62718

12	1492	81.0	560	22	AAB62724	Borrelia sp chimera
13	1490	80.9	392	22	AAB62734	Borrelia sp chimera
14	1488	80.8	384	22	AAB62726	Borrelia sp chimera
15	1485	80.6	408	22	AAB62737	Borrelia sp chimera
16	1284.5	69.7	369	22	AAB62716	Borrelia sp chimera
17	1281.5	69.6	363	22	AAB62731	Borrelia sp chimera
18	1248.5	67.8	367	22	AAB62719	Borrelia sp chimera
19	1245.5	67.6	391	22	AAB62735	Borrelia sp chimera
20	1213.5	65.9	378	22	AAB62725	Borrelia sp chimera
21	1212.5	65.8	369	22	AAB62714	Borrelia sp chimera
22	1209.5	65.7	393	22	AAB62732	Borrelia sp chimera
23	1205.5	65.4	401	22	AAB62733	Borrelia sp chimera
24	1181.5	64.1	368	22	AAB62717	Borrelia sp chimera
25	1173.5	63.7	391	22	AAB62736	Borrelia sp chimera
26	1165	63.2	410	22	AAB62740	Borrelia sp chimera
27	1164	63.2	386	22	AAB62727	Borrelia sp chimera
28	964	52.3	466	16	AAR5740	B31 outer surface
29	928	50.4	192	22	AAB62721	B burgdorferi ospc
30	928	50.4	210	16	AAW11935	B. burgdorferi prot
31	928	50.4	210	16	AAR75727	B. burgdorferi str
32	927	50.3	587	16	AAR75746	B31 Osp-A/antigen
33	926	50.3	192	15	AAR60886	Borrelia sp2 ospc
34	920	49.9	209	22	AAB62720	B burgdorferi ospc
35	915	49.7	466	16	AAR75739	B31 outer surface
36	908	49.3	192	22	AAB62703	B burgdorferi ospc
37	885	48.0	190	22	AAB62707	B burgdorferi ospc
38	879	47.7	192	15	AAR60889	Borrelia sp2 antlg
39	864	46.9	176	15	AAR62772	Borrelia B31 antlg
40	852	46.3	176	15	AAR62775	Borrelia 297 antlg
41	694.5	37.7	177	15	AAR62773	Borrelia 25015 ant
42	694.5	37.7	177	15	AAR60887	Borrelia 25015 osp
43	678.5	36.8	193	22	AAB62709	B burgdorferi ospc
44	676	36.7	212	16	AAW11934	B outer surface prot
45	674.5	36.6	191	15	AAR60897	Borrelia VS461 ant

#### ALIGNMENTS

RESULT 1  
ID AAB62710 standard; Protein: 374 AA.  
XX  
AC AAB62710;  
XX  
XX 03-APR-2001 (first entry)  
XX  
DE Borrelia sp chimeric ospc protein seq ID NO: 24.  
XX  
KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
OS Chimeric - Borrelia sp.  
XX  
OS Chimeric - Borrelia sp.  
XX  
PN WO200078966-A1.  
XX  
XX 28-DEC-2000.  
XX  
XX 19-JUN-2000; 2000MO-US16915.  
XX  
XX 18-JUN-1999; 99US-0140042.  
XX  
XX (UYNV ) UNIV NEW YORK STATE RES FOUND.  
XX  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX  
XX WPI; 2001-050113/05.  
XX  
XX N-PSDB; AAF29014.  
XX  
XX Compositions of ospc polypeptides from strains of Borrelia which cause  
XX  
XX Lyme disease are used to immunize animals and detect immune responses  
XX  
XX to Lyme disease -

XX Claim 43; Page 78-79; 160pp; English.  
 PS The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 374 AA;

Query Match  
 Best Local Similarity 99.7%; Score 1837; DB 22; Length 374;  
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACNNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 61  
 2 AKNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 60  
 3 GKKIHQNNGLDTEYNHNGSLAGAVAIITLIKOKLDGLKNGEIKEDDAKCKSEFTTNK 121  
 4 61 gkkihngngldteynhngslagavaiistlikgldgkngelkexidaackseftnk 120  
 5 122 LKEKHTDLGEGYTDADAEALIKTNGTKGAEIGKTESVEVLSKAKEMIANSVKE 181  
 6 121 lkehhdldggytdadakealiktngtkgaeigktesvevlskaakemiansvke 180  
 7 182 LTPVVAESPAMVNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 241  
 8 181 ltpvvaespamvnsqkqntsansadesvkgpmlteiskitdsnavlavkeveall 240  
 9 242 TSIDEIAKAIGKIKINDVSLDNEADHNGSLISGAVLISNLTIRKTSATKDSGEIAKAEIER 301  
 10 241 tsidelakaigkikndvslndeadhngslisgavlisnltirktSATKDSGEIAKAEIER 300  
 11 302 AKKSEEFATKIGEGYTDGEGYTDNNAKKAILETNDKTKGAEIGKTESVEVLSKAKEMIANSVKE 361  
 12 301 akkseefatkigegtytdggytdnnaakkailetnDKTKGAEIGKTESVEVLSKAKEMIANSVKE 360  
 13 362 AKEMLTNSVKEIUS 375  
 14 361 akemltnsvkeiUS 374

## RESULT 2

62728  
 AAB62728 standard; Protein; 398 AA.  
 AAB62728;

03-APR-2001 (first entry)

Borrelia sp chimeric ospc protein SEQ ID NO: 60.

Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 XX Chimeric - Borrelia sp.  
 PN WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.

(UNNY) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOLOGICALS INC.

Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

DR WPI: 2001-050113/06.  
 DR N-PSDE: AAF29032.  
 XX Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease.  
 XX  
 PS Claim 43; Page 123; 160pp; English.  
 PS The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 398 AA;

Query Match  
 Best Local Similarity 99.6%; Score 1834; DB 22; Length 398;  
 Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ACNNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 61  
 2 AKNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 84  
 3 GKKIHQNNGLDTEYNHNGSLAGAVAIITLIKOKLDGLKNGEIKEDDAKCKSEFTTNK 121  
 4 85 gkkihngngldteynhngslagavaiistlikgldgkngelkexidaackseftnk 144  
 5 122 LKEKHTDLGEGYTDADAEALIKTNGTKGAEIGKTESVEVLSKAKEMIANSVKE 181  
 6 121 lkehhdldggytdadakealiktngtkgaeigktesvevlskaakemiansvke 180  
 7 182 LTPVVAESPAMVNSGKQNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 241  
 8 181 ltpvvaespamvnsqkqntsansadesvkgpmlteiskitdsnavlavkeveall 240  
 9 242 TSIDEIAKAIGKIKINDVSLDNEADHNGSLISGAVLISNLTIRKTSATKDSGEIAKAEIER 301  
 10 241 tsidelakaigkikndvslndeadhngslisgavlisnltirktSATKDSGEIAKAEIER 300  
 11 302 AKKSEEFATKIGEGYTDGEGYTDNNAKKAILETNDKTKGAEIGKTESVEVLSKAKEMIANSVKE 361  
 12 325 akkseefatkigegtytdggytdnnaakkailetnDKTKGAEIGKTESVEVLSKAKEMIANSVKE 360  
 13 362 AKEMLTNSVKEIUS 375  
 14 385 akemltnsvkeiUS 398

## RESULT 3

62738  
 AAB62738 standard; Protein; 401 AA.  
 AAB62738;

03-APR-2001 (first entry)

Borrelia sp chimeric ospc protein SEQ ID NO: 80.

Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 XX Chimeric - Borrelia sp.  
 PN WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.



XX (UNY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
XX Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
DR WPI: 2001-050113/06.  
N-PSDB: AAF29042.  
XX  
XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -  
XX  
XX Claim 43; Page 150-151; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX  
XX Sequence 401 AA:  
Query Match 98.8%; Score 1819.5; DB 22; Length 401;  
Best Local Similarity 98.7%; Pred. No. 5.6e-118;  
Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;  
XX  
OY 2 ACNNSGKDCNTSANSADSEVKGPNLTETISKRITDSNAVLLAVEFALLSSIDEIAKAI 61  
DB :|||||  
DB 25 scnngskdgnntsaansadesvkgpnltetiskritdsnavllavefalleallssideiaakal 84  
OY 62 GKRIHONNGLDTEYHNHNSGLAGAVAIISPLIKOKLDGLNKEGKEKIDAAKCCSEFTTNK 121  
DB 85 gkllhngldteyhnngslagayaistlikgldglnkegkidaakkcsefttnk 144  
OY 122 LKEKHTDLCKEGVTDADAEALIKTNGTKGAEELGKLFESYEVUSKAKEMLANSVK 181  
DB :|||||  
DB 145 lkekhndlgkegvtddaealiktngtkgaeelgklfesyeviskaakemlansvke 204  
OY 182 LKSPVVAES---PAMVNNNGKDCNTSANSADSEVKGPNLTETISKRITDSNAVLLAVEF 238  
DB :|||||  
DB 205 lksppvaesppkpsmwnsgkdnntsaansadesvkgpnltetiskritdsnavllaveve 264  
OY 239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTIRKISAIKDSGELKAE 298  
DB :|||||  
OY 265 elltsidelakelgkllkndvslndeadhngslisgayllsnlttkksaiksgegelkae 324  
OY 299 TEKAKKCEEFPAKLKGEFTDLGKEGVTDDNAKKAAILKTNNDKTKGADELEKLFESYKVL 358  
DB :|||||  
OY 325 ltkakkcseeflaklgehtldlgkegvtddnakkalktndktgadelkelfesvkn 384  
OY 359 SKAAKEMLTNSVKEKLS 375  
DB :|||||  
DB 385 skaakemltnsvkeKLS 401

RESULT 4  
AAB62712  
ID AAB62712 standard; Protein: 378 AA.  
XX  
XX AAB62712;  
XX  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospc protein SEQ ID NO: 28.  
XX  
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
XX Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX  
XX WO200078966-A1.

XX 28-DEC-2000.  
PD  
XX 19-JUN-2000; 2000MO-US16915.  
PE  
XX 18-JUN-1999; 99US-0140042.  
PR  
XX  
XX (UNY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
XX Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
DR WPI: 2001-050113/06.  
N-PSDB: AAF29016.  
XX  
XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -  
XX  
XX Claim 43; Page 83-84; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX  
XX Sequence 378 AA:  
Query Match 98.3%; Score 1810.5; DB 22; Length 378;  
Best Local Similarity 98.1%; Pred. No. 2.2e-117;  
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;  
XX  
OY 1 MACNNSGKDCNTSANSADSEVKGPNLTETISKRITDSNAVLLAVEFALLSSIDEIAKA 60  
DB :|||||  
DB 1 macnngskdgnntsaansadesvkgpnltetiskritdsnavllavefalleallssideiaaka 60  
OY 61 IGKRIHONNGLDTEYHNHNSGLAGAVAIISPLIKOKLDGLNKEGKEKIDAAKCCSEFTTN 120  
DB :|||||  
DB 61 lggkllhngldteyhnngslagayaistlikgldglnkegkidaakkcsefttn 120  
OY 121 KLEKHTDLCKEGVTDADAEALIKTNGTKGAEELGKLFESYEVUSKAKEMLANSVK 180  
DB :|||||  
DB 121 klekhndlgkegvtddaealiktngtkgaeelgklfesyeviskaakemlansvke 180  
OY 181 ELKSPVVAES---PAMVNNNGKDCNTSANSADSEVKGPNLTETISKRITDSNAVLLAVEF 237  
DB :|||||  
DB 181 elksppvaesppkpsmwnsgkdnntsaansadesvkgpnltetiskritdsnavllavev 240  
OY 238 ETLTSLIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTIRKISAIKDSGELKA 297  
DB :|||||  
DB 241 elltsidelakelgkllkndvslndeadhngslisgayllsnlttkksaiksgegelka 300  
OY 298 ETEKAKKCEEFPAKLKGEFTDLGKEGVTDDNAKKAAILKTNNDKTKGADELEKLFESYK 357  
DB :|||||  
DB 301 etekakkcseeflaklgehtldlgkegvtddnakkalktndktgadelkelfesvkn 360  
OY 358 LSKAAKEMLTNSVKEKLS 375  
DB :|||||  
DB 361 lskaaakemltnsvkeKLS 378

RESULT 5  
AAB62711  
ID AAB62711 standard; Protein: 373 AA.  
XX  
XX AAB62711;  
XX  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospc protein SEQ ID NO: 26.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 WPI: 2001-050113/06.  
 N-PSDB: AAF29015.  
 XX  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 XX to Lyme disease -  
 PS Claim 43; Page 81; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 373 AA:

Query Match 86.5%; Score 1592.5; DB 22; Length 373;  
 Best Local Similarity 87.7%; Pred. No. 2.4e-102;  
 Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;  
 QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 61  
 DB 1 acnnsqkdgntsansadesvkgpnlteiskrtidsnavllavkeveallssideiaakai 60  
 QY 62 GKRIHONNGLDTEYNNHNSLAGAVAISTLIKQKLDGLKNEGKKEKIDAAKCEFTFNK 121  
 DB 61 gkrihngngldteyhnngslagayaistlikqkldglnegkkekidaakcsetftnk 120  
 QY 122 LKEKHTDLGKEGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181  
 DB 121 lkehtldlgkegvtdadaekaelikngtktgaeelgklfesvevlskaakemlansvke 180  
 QY 182 LTPVVAESPAMVNNSGDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVETTL 241  
 DB 181 ltpvvaespamvnnsgdgntsansadesvkgpnlteiskrtidsnavllavkeletll 240  
 QY 242 TSIDELA-KAIGKRIKNDVSLDNEADHNSLISGAYVLSNLTIKKISAIKDSSELAETIE 300  
 DB 241 asidelatkaigkriqngslagaveagngcllagayvlsnltikgldglnsekilekie 300  
 QY 301 KAKCSEFTAKLGEHTDLGKEGVTDADAKAELIKTNDKTGADDELKLFESVKNLSK 360  
 DB 301 nakcseftaklgehngldgkegvtdadakaellitdaakdgaelelrfkavenlak 360  
 QY 361 AAKEMLTNSVKEEL 373  
 DB 361 aakemlansvkel 373

RESULT 6  
 AAB62729  
 ID AAB62729 standard; Protein; 397 AA.  
 XX

AC AAB62729;  
 XX  
 DT 03-APR-2001 (First entry)  
 XX  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 62.  
 XX  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 WPI: 2001-050113/06.  
 N-PSDB: AAF29033.  
 XX  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 XX to Lyme disease -  
 PS Claim 43; Page 125-126; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 397 AA:

Query Match 86.3%; Score 1589.5; DB 22; Length 397;  
 Best Local Similarity 87.4%; Pred. No. 4.2e-102;  
 Matches 326; Conservative 21; Mismatches 23; Indels 1; Gaps 1;  
 QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 61  
 DB 25 scnnsqkdgntsansadesvkgpnlteiskrtidsnavllavkeveallssideiaakai 84  
 QY 62 GKRIHONNGLDTEYNNHNSLAGAVAISTLIKQKLDGLKNEGKKEKIDAAKCEFTFNK 121  
 DB 85 gkrihngngldteyhnngslagayaistlikqkldglnegkkekidaakcsetftnk 144  
 QY 122 LKEKHTDLGKEGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181  
 DB 145 lkehtldlgkegvtdadaekaelikngtktgaeelgklfesvevlskaakemlansvke 204  
 QY 182 LTPVVAESPAMVNNSGDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVETTL 241  
 DB 205 ltpvvaespamvnnsgdgntsansadesvkgpnlteiskrtidsnavllavkeletll 264  
 QY 242 TSIDELA-KAIGKRIKNDVSLDNEADHNSLISGAYVLSNLTIKKISAIKDSSELAETIE 300  
 DB 265 asidelatkaigkriqngslagaveagngcllagayvlsnltikgldglnsekilekie 324  
 QY 301 KAKCSEFTAKLGEHTDLGKEGVTDADAKAELIKTNDKTGADDELKLFESVKNLSK 360  
 DB 325 nakcseftaklgehngldgkegvtdadakaellitdaakdgaelelrfkavenlak 384  
 QY 361 AAKEMLTNSVKEEL 373  
 DB 385 aakemlansvkel 397

```

RESULT 7
AAB62739
ID AAB62739 standard; Protein; 400 AA.
XX
AC AAB62739;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.
XX
KV Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000; 2000WO-US16915.
XX
PA (UNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29043.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 153; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 400 AA;

Query Match 85.5%; Score 1575; DB 22; Length 400;
Best Local Similarity 86.4%; Pred. No. 4.3e-101;
Matches 325; Conservative 22; Mismatches 25; Indels 4; Gaps 2;

```

```

DB 325 kienakcsedftkklegenhqgienvdenakkallitdaakdgaaeilekIfkaven 384
QY 358 LSKAKEMLTNSVKEL 373
DB 385 lakaakemlansvkel 400

RESULT 8
AAB62713
ID AAB62713 standard; Protein; 377 AA.
XX
AC AAB62713;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 30.
XX
KV Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000; 2000WO-US16915.
XX
PA (UNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29017.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 86-87; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 377 AA;

Query Match 85.0%; Score 1566; DB 22; Length 377;
Best Local Similarity 85.9%; Pred. No. 1.7e-100;
Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

```

```

Qy 238 ETLTSLIDELA-KAIGKRKINDSLDNEADHNSGLISGAYLISLITRKISAITDSEGLK 296
Db 241 ETLTSLIDELAETKAIKGIQNGGLAVEAGINGLIAGYATISKLIITQKIDGLKNSK 300
Qy 297 AEIERAKKCSSEFTAKLTGHEHTDGLKEGVTDDNARKKAILKLTNNDKTKGADELEKLFESVK 356
Db 301 EKLTKAKCCSEDFLKKIIEGHAQYIENVDENAKKAILLITDAKDKGAAELEKLFIVE 360
Qy 357 NLKRAKEMLTNSYKEL 373
Db 361 nlakaakemlansvkel 377

RESULT 9
AAB62715
ID AAB62715 standard; Protein; 370 AA.

AAB62715:
03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 34.
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
XX Chimeric - Borrelia sp.
XX WO200078966-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UNYV) UNIV NEW YORK STATE RES FOUND.
XX (BROO) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BU, Gomes-Solecki M;
XX WPI. 2001-050113/06.
XX DR N-PSDB; AAF29019.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 91-92; 160pp; English.
XX
XX PS The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX
XX Sequence 370 AA:

Query Match 83.0%; Score 1529; DB 22; Length 370;
Best Local Similarity 87.0%; Pred. No. 5,8e-98;
Matches 327; Conservative 9; Mismatches 32; Indels 8; Gaps 4;
Qy 2 AANNNGSGKGNSTANSADSVKRPNTLEISKTTSDNSAVILAKVEYBALISIDETAAKAI 61
Db 1 acnnsqkdgntsaansadesvkgpnltelskkltdsnavllavkeveallssidel-akaI 59
Qy 62 GKRIHQNNGLDIEYVHNHNSGLAGAYASTPLIKOKIDGK-NGELKEKIDAAKCSSEFPFN 120
Db 60 gkllndsgldneannnslagaytistlltqklskngseglleklaaakcssefst 119
Qy 121 KLKEKHTDLGEGVTADDAKKAAILTKNGT-KTKGAEEIGKLFEESEVYLSKAKEMLANSV 179

```

```

Db      120 klnhdnhaqlqdygtlndekkaiklkanaagdkdysveelksgslstskaaeklansv 179
QY      180 KELTSPVVAASPAWVNNSGKDGNTSANSADSVKGPMLTEISKRTITESNAVLAVKEVEET 239
Db      180 keltsprvhg-----nsngkdgnstsansadesvkgpmltelskkltessnavlavkevet 234
QY      240 LNTSIDELAIAIGKKIKNDVSLDNEADHNGSLISGAVLISLTIRKITSATRKDSGLKAEI 299
Db      235 llsldelaikaigklkndvsldeahngslisgavylisslltkkksalkdsgeklaei 294
QY      300 EKAKKCSSEFTAKLKGEHTDLGREGVTDDNAAKAILTNNDKTRKGADELKLEFESVKNL 359
Db      295 ekakkcsseftaklkgshtdlkgreyrddnakkailtntndktkgadeleklfesvnl 354
QY      360 KAAKEMITNSVKELTS 375
Db      355 kaakemltinsvkelts 370

RESULT 10
AAB62730
ID      AAB62730 standard; Protein; 394 AA.
AC      AAB62730;
DT      03-APR-2001 (first entry)
DE      Borrelia sp chimeric ospc protein SEQ ID NO: 64.
KW      Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX      OS Chimeric - Borrelia sp.
XX      OS Chimeric - Borrelia sp.
XX      PN WO200078966-A1.
XX      PD 28-DEC-2000.
XX      PF 19-JUN-2000; 2000WO-US16915.
XX      PR 18-JUN-1999; 99US-0140042.
XX      PA (UUYV ) UNIV NEW YORK STATE RES FOUND.
XX      PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX      PI Datwyler RJ, Seinst G, Dykhuzen D, Luft BJ, Gomes-Solecki M;
XX      DR WPI: 2001-050113/06.
XX      N-PSDB; AAF29034.
PS      Compositions of OspC polypeptides from strains of Borrelia which cause
PT      Lyme disease are used to immunize animals and detect immune responses
PR      to Lyme disease -
XX      PS Claim 43; Page 128-129; 160pp; English.
CC      The present invention provides compositions comprising ospc proteins and
CC      chimeric ospc proteins from members of the Borrelia genus. These may be
CC      Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC      vaccines against Borrelia infection, which is spread by ticks and leads
CC      to Lyme disease.
XX      XX Sequence 394 AA;
SQ
Query Match 82.8%; Score 1526; DB 22; Length 394;
Best Local Similarity 86.7%; Pred. No. 1e-97;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

```

QY 62 GKRIHONNGLDTEYHNHNSGLAGAVASTLIKOKLDGLK-NEGCKEKIDAAKCKSETFTN 120  
 Db 84 gkkinndgsldneanrneesllagaylslcltqklskngseglakeiaakckseefst 143  
 QY 121 KKEKHNDGKRGVNDADAKKALIKTNGT-KTGAEELGKLFESVYLSAAKEMLANSV 179  
 Db 144 ktkdnhaqiglgvtdenakkaillkanaagkdkgyveelelsglseslakaakemlanav 203  
 QY 180 KELTSPVVAESPAMVNSGKDGNTSANSADSEYKGNLTETSKITESNAVLAVERET 239  
 Db 204 keltepvng-----nsgkdgnltsansadesvkgpnlelskkltesnavlaverev 258  
 QY 240 LTSTDELAKAIGKRIKNDVSLDNEADHNSGLISGAYLISNLTITKRIKISAKDSGELKAET 299  
 Db 259 lltstidelakalqkklkndvslidneadngslisgayllsnlltkklsalksdsgelelkael 318  
 QY 300 EKAKCKSEEFKAKGHEHTDLGEGVTDNNAKKAIIKTNNDKTKGADLEKLFESYKNTLS 359  
 Db 319 ekakckseefkalkyehcdlggyvtdnakkalltkmndckgadelkltesyknls 378  
 Db 360 KAKEMLTNSVKELTS 375  
 Db 379 kaekmltnsvkelts 394  
 RESULT 11  
 AAB62718  
 ID AAB62718 standard; Protein: 368 AA.  
 AC AAB62718;  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 40.  
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric; protein; tick.  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 PD 28-DEC-2000.  
 PF 19-JUN-2000; 2000WO-US16915.  
 PR 18-JUN-1999; 99US-0140042.  
 XX (UYNY ) UNITV NEW YORK STATE RES FOUND.  
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Seinoat G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR MPI: 2001-050113/06.  
 DR N-PSDB: AAF29022.  
 PT Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 PS Claim 43; Page 99-100; 160pp; English.  
 XX The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX Sequence 368 AA;  
 SQ  
 Query Match 81.1%; Score 1493; DB 22; Length 368;

Best Local Similarity 83.4%; Pred. No. 1.8e-95;  
 Matches 312; Conservative 22; Mismatches 34; Indels 6; Gaps 2;  
 QY 2 ACNNSGKDGNTSANSADSEYKGNLTETSKITESNAVLAVEREALLSSIDEIAAKAI 61  
 Db 1 acnnsygdgnasansadesvkgpnltetsskltesnavlaverevcllasidelakeal 60  
 QY 62 GKRIHONNGLDTEYHNHNSGLAGAVASTLIKOKLDGLKNEGCKEKIDAAKCKSETFTNK 121  
 Db 61 gkkl-ngnleangskntslisgaysldlaeklnvlnkeelkeldtakqscsteflnk 119  
 QY 122 LKEHNTDLGKRGVNDADAKKALIKTNGTGTGAEELGKLFESVYLSAAKEMLANSVKE 181  
 Db 120 lkeehavlgldnlddnagralikhankdkgyaelelklfavenlskaagclknayke 179  
 QY 182 LTSPVVAESPAMVNSGKDGNTSANSADSEYKGNLTETSKITESNAVLAVERETLL 241  
 Db 180 ltsplvng-----nsgkdgnltsansadesvkgpnlelskkltesnavlaverevll 234  
 QY 242 TSIDELAKAIGKRIKNDVSLDNEADHNSGLISGAYLISNLTITKRIKISAKDSGELKAET 301  
 Db 235 tsidelakalqkklkndvslidneadngslisgayllsnlltkklsalksdsgelelkael 294  
 QY 302 AKKCKSEEFKAKGHEHTDLGEGVTDNNAKKAIIKTNNDKTKGADLEKLFESYKNTLSKA 361  
 Db 295 akckseefkalkyehcdlggyvtdnakkalltkmndckgadelkltesyknlska 354  
 QY 362 AKEMLTNSVKELTS 375  
 Db 355 ekemltnsvkelts 368  
 RESULT 12  
 AAB62724  
 ID AAB62724 standard; Protein: 560 AA.  
 AC AAB62724;  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 52.  
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric; protein; tick.  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 PD 28-DEC-2000.  
 PF 19-JUN-2000; 2000WO-US16915.  
 PR 18-JUN-1999; 99US-0140042.  
 XX (UYNY ) UNITV NEW YORK STATE RES FOUND.  
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Seinoat G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR MPI: 2001-050113/06.  
 DR N-PSDB: AAF29028.  
 PT Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 PS Claim 43; Page 112-113; 160pp; English.  
 XX The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 560 AA;

Query Match 81.0%; Score 1492; DB 22; Length 560;  
 Best Local Similarity 83.5%; Pred. No. 3.6e-95;  
 Matches 313; Conservative 20; Mismatches 40; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLETISKITDSNAVLAVKEVEALLSIDEIAAKA 60  
 DB 1 macnsgkdqntsaansadesvkgnltelskltidnavlavkeveallsideiaaka 60  
 QY 61 IGGRIHONNGLDTEYVNHNSLAGAVASTLIKOKLDGKNEGKLEKIDAAKCEFTFN 120  
 DB 61 iggkhhnngldteyhnngslagayaistlikqkldgkneglekidaakcseftfn 120  
 QY 121 KLEKHTDLGKEGVTADAKKAIKTNGTKTGAEELKLESEVVLSSKAAKEMLANSVK 180  
 DB 121 klekhthdlgkegvtdadakeaiktngtktgaeelgklesvevlskaakemlansvk 180  
 QY 181 ELTSPVAESPAMVNSGKDGNT-SANSADSVKGPNTLETISKITTESNAVLAVKEVEET 239  
 DB 181 eltspvaaespamvnsqkdgdsastnpadesakgnltelskltidnavlavkevet 240  
 QY 240 LITSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTKKISAKDSGELKAE 298  
 DB 240 litstidelakkaigkikndvslneadhnslisgayslnttkkisaikdsgelekae 298  
 QY 299 IEKAKKSEETAKLKGEHTDLGKEGVTDDNAKKAILETNNDKTGADDELKLESEVKNL 358  
 DB 301 iekakkeetfaklkgehtdlgksgvadtdnahkaailktnndktgadelklesvkn 360  
 QY 359 SKAAKEMLTNSVKEL 373  
 DB 361 lkaaqvaltnsvkel 375

RESULT 13

AAB62734 standard; Protein; 392 AA.

AC AAB62734;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX

Borrelia sp chimeric ospc protein SEQ ID NO: 72.

Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.

(UANY ) UNIV NEW YORK STATE RES FOUND.

(BROO-) BROOK BIOTECHNOLOGIES INC.

Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.

N-PSDB; AAF29038.

Compositions of ospc polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

PS Claim 43; Page 139-140; 160pp; English.

XX The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX

SQ Sequence 392 AA;

Query Match 80.9%; Score 1490; DB 22; Length 392;  
 Best Local Similarity 83.2%; Pred. No. 3.1e-95;  
 Matches 311; Conservative 23; Mismatches 34; Indels 6; Gaps 2;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLETISKITDSNAVLAVKEVEALLSIDEIAAKA 61  
 DB 25 scnnsqkdqnsaansadesvkgnltelskltidnavlavkevetllsideiatkai 84  
 QY 62 GKRIHONNGLDTEYVNHNSLAGAVASTLIKOKLDGKNEGKLEKIDAAKCEFTFN 121  
 DB 85 gkkihngldteyhnngslagayaistlikqkldgkneglekidaakcseftfn 143  
 QY 122 LKEKHTDLGKEGVTADAKKAIKTNGTKTGAEELKLESEVVLSSKAAKEMLANSVK 181  
 DB 144 lkekhthdlgkegvtdadakeaiktngtktgaeelgklesvevlskaakemlansvk 203  
 QY 182 ELTSPVAESPAMVNSGKDGNTSANSADSVKGPNTLETISKITTESNAVLAVKEVEETL 241  
 DB 204 eltspvaaespamvnsqkdgdsastnpadesakgnltelskltidnavlavkevetll 258  
 QY 242 TSIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTKKISAKDSGELKAE 301  
 DB 259 tsidelakkaigkikndvslneadhnslisgayslnttkkisaikdsgelekae 318  
 QY 302 AKKCEETAKLKGEHTDLGKEGVTDDNAKKAILETNNDKTGADDELKLESEVKNL 361  
 DB 319 akkceetfaklkgehtdlgksgvadtdnahkaailktnndktgadelklesvkn 378  
 QY 362 AKEMLTNSVKEL 375  
 DB 379 akemltnsvkel 392

RESULT 14

AAB62726 standard; Protein; 384 AA.

AC AAB62726;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX

Borrelia sp chimeric ospc protein SEQ ID NO: 56.

Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.

(UANY ) UNIV NEW YORK STATE RES FOUND.

(BROO-) BROOK BIOTECHNOLOGIES INC.

Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.

DR N-PSDB; AAF29030.  
 XX Compositions of OspC polypeptides from strains of *Borrelia* which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 117-118; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the *Borrelia* genus. These may be  
 CC *Borrelia burgdorferi*, *B. afzelii* or *B. garinii*. These can be used as  
 CC vaccines against *Borrelia* infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 XX Sequence 384 AA:  
 SQ  
 Query Match 80.8%; Score 1488; DB 22; Length 384;  
 Best Local Similarity 82.0%; Pred. No. 4.1e-95;  
 Matches 309; Conservative 31; Mismatches 31; Indels 6; Gaps 3;  
 DB 1 ACNNSGKDGNTSANSADSEYKGNLTREISKRTDSNAVLLAVKEVALLSSIDEIAAKAI 61  
 1 acnnsqkdgntnsadsesvkgpnltreiskrtidsnavllavkevallssideiaakal 60  
 QY 62 GKRIHONNGIDTEYNHNGSLLAGAVAIISTLIKOKLDKNEGLKEKRIIDAKKCEFTFNK 121  
 61 gkrihngngldteynhngslilagayaisllkqkloglkneglkeikidaakkcseftfnk 120  
 QY 122 LKEKHHTDLGEGYTDADAKAEILKTNGTKGAELGKLFESYEVLSKAKEMLANSVKE 181  
 121 lkekhhtdlggytdadakeailktngtkgaeelgklfesvevlskaakemlansvke 180  
 QY 182 LTPVVAESPAMV--NNSGKDGNTSANSADSEYKGNLTREISKRTTESNAVLLAVKEVE 238  
 181 ltpvvaespampkpfhngnsq--gdsasnpdesakgpnltreiskrtidsnavllavkeve 238  
 DB 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKRIKISAKSGELKAE 298  
 239 allssideiskalgkrikndvsl dneadnhngslisgaylislntitkrisaksgelk 297  
 QY 299 IERAKKCEEFATAKLGHEHTDLGEGYTDNNAKKAIIKTNNDKTKGADELEKLFESYKVL 358  
 298 lkeakdcsgkfttklkdshaeligsvqdnaakkaillktngtkdgakeleelfkstlesl 357  
 QY 359 SKAAKEMLTNSVKELTS 375  
 358 skaagaaltnsvkeltn 374  
 DB 358 skaagaaltnsvkeltn 374  
 UTR 15  
 AAB62737  
 ID AAB62737 standard; Protein; 408 AA.  
 AC AAB62737;  
 XX  
 XX 03-APR-2001 (first entry)  
 DE *Borrelia* sp chimeric ospc protein SEQ ID NO: 78.  
 XX  
 XX *Borrelia*; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 KM  
 XX Chimeric - *Borrelia* sp.  
 OS Chimeric - *Borrelia* sp.  
 XX  
 XX WO200078966-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX 19-JUN-2000; 2000WO-US16915.  
 PF  
 XX 18-JUN-1999; 99US-0140042.  
 PR  
 XX

PA (UNYV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 XX Datkwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PSDB; AAF29041.  
 XX  
 PS Claim 43; Page 147-148; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the *Borrelia* genus. These may be  
 CC *Borrelia burgdorferi*, *B. afzelii* or *B. garinii*. These can be used as  
 CC vaccines against *Borrelia* infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 XX Sequence 408 AA:  
 SQ  
 Query Match 80.6%; Score 1485; DB 22; Length 408;  
 Best Local Similarity 81.7%; Pred. No. 7.2e-95;  
 Matches 308; Conservative 32; Mismatches 31; Indels 6; Gaps 3;  
 DB 2 ACNNSGKDGNTSANSADSEYKGNLTREISKRTDSNAVLLAVKEVALLSSIDEIAAKAI 61  
 2 acnnsqkdgntnsadsesvkgpnltreiskrtidsnavllavkevallssideiaakal 60  
 QY 25 SCNSGKDGNTSANSADSEYKGNLTREISKRTIDSNAVLLAVKEVALLSSIDEIAAKAI 84  
 25 scnsgkdgntnsadsesvkgpnltreiskrtidsnavllavkevallssideiaakal 84  
 QY 62 GKRIHONNGIDTEYNHNGSLLAGAVAIISTLIKOKLDKNEGLKEKRIIDAKKCEFTFNK 121  
 61 gkrihngngldteynhngslilagayaisllkqkloglkneglkeikidaakkcseftfnk 120  
 DB 122 LKEKHHTDLGEGYTDADAKAEILKTNGTKGAELGKLFESYEVLSKAKEMLANSVKE 181  
 121 lkekhhtdlggytdadakeailktngtkgaeelgklfesvevlskaakemlansvke 180  
 QY 182 LTPVVAESPAMV--NNSGKDGNTSANSADSEYKGNLTREISKRTTESNAVLLAVKEVE 238  
 181 ltpvvaespampkpfhngnsq--gdsasnpdesakgpnltreiskrtidsnavllavkeve 238  
 DB 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKRIKISAKSGELKAE 298  
 239 allssideiskalgkrikndvsl dneadnhngslisgaylislntitkrisaksgelk 297  
 QY 299 IERAKKCEEFATAKLGHEHTDLGEGYTDNNAKKAIIKTNNDKTKGADELEKLFESYKVL 358  
 298 lkeakdcsgkfttklkdshaeligsvqdnaakkaillktngtkdgakeleelfkstlesl 357  
 QY 359 SKAAKEMLTNSVKELTS 375  
 358 skaagaaltnsvkeltn 398  
 DB 382 skaagaaltnsvkeltn 398

Search completed: March 18, 2002, 09:54:29  
 Job time: 327 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:47 ; Search time 55.5 seconds  
(without alignments)  
152.049 Million cell updates/sec

Title: US-09-596-746A-24

Perfect score: 1842  
Sequence: 1 MACNNSGKCDNTSANSADSE.....KNLSKAKEMUTNSVKELTS 375

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	52.3	466	4	US-08-235-836C-110 Sequence 110, Appl
2	932	50.6	210	4	US-08-235-836C-30 Sequence 30, Appl
3	928	50.4	209	4	US-09-196-293-15 Sequence 15, Appl
4	928	50.4	210	1	US-08-158-353-3 Sequence 3, Appl
5	927	50.3	588	4	US-08-235-836C-122 Sequence 122, Appl
6	921	50.0	210	4	US-08-209-603E-15 Sequence 15, Appl
7	915	49.7	466	4	US-08-235-836C-107 Sequence 107, Appl
8	676	36.7	212	1	US-08-158-353-2 Sequence 2, Appl
9	661	35.9	212	1	US-08-031-295-2 Sequence 2, Appl
10	661	35.9	212	1	US-07-903-580-2 Sequence 2, Appl
11	614	33.3	212	1	US-08-158-353-4 Sequence 4, Appl
12	612	33.2	212	4	US-09-196-293-11 Sequence 11, Appl
13	612	33.2	212	4	US-08-209-603E-11 Sequence 11, Appl
14	609	33.1	212	4	US-08-235-836C-34 Sequence 34, Appl
15	600	32.6	209	4	US-08-235-836C-32 Sequence 32, Appl
16	583.5	31.7	207	4	US-08-235-836C-36 Sequence 36, Appl
17	163.5	8.9	2482	1	US-08-328-234-6 Sequence 6, Appl
18	162.5	8.8	3248	5	US-08-353-700-1 Sequence 1, Appl
19	162.5	8.8	3248	5	PCT-US95-16216-1 Sequence 1, Appl
20	157	8.5	194	4	US-09-364-083-2 Sequence 2, Appl
21	148.5	8.1	1713	3	US-08-600-982-24 Sequence 24, Appl
22	148.5	8.1	1713	5	PCT-US94-10261A-24 Sequence 24, Appl
23	147	8.0	1164	4	US-08-923-992A-2 Sequence 2, Appl
24	146	7.9	1196	1	US-08-144-121-4 Sequence 4, Appl
25	146	7.9	1196	2	US-08-735-893-4 Sequence 4, Appl
26	145.5	7.9	1388	2	US-08-685-576-1 Sequence 1, Appl
27	145	7.9	1588	5	PCT-US93-07261-11 Sequence 11, Appl

28	145	7.9	1663	5	PCT-US93-07261-16 Sequence 16, Appl
29	145	7.9	3111	2	US-08-460-309-4 Sequence 4, Appl
30	145	7.9	3111	2	US-08-125-077-4 Sequence 4, Appl
31	143.5	7.8	630	4	US-08-973-462-9 Sequence 9, Appl
32	143	7.8	1164	4	US-08-923-992A-10 Sequence 10, Appl
33	142.5	7.7	688	3	US-09-141-047-8 Sequence 8, Appl
34	141	7.7	1786	4	US-08-973-462-8 Sequence 8, Appl
35	139	7.5	2285	4	US-09-308-375-2 Sequence 2, Appl
36	138.5	7.5	1098	4	US-08-923-992A-8 Sequence 8, Appl
37	138	7.5	1128	4	US-08-923-992A-6 Sequence 6, Appl
38	137	7.4	1104	4	US-08-923-992A-4 Sequence 4, Appl
39	136	7.4	1388	2	US-08-685-576-4 Sequence 4, Appl
40	135.5	7.4	573	4	US-08-235-836C-112 Sequence 112, Appl
41	133.5	7.2	641	4	US-08-961-083-160 Sequence 160, Appl
42	132.5	7.2	1561	3	US-08-894-017-23 Sequence 23, Appl
43	131.5	7.1	1068	4	US-09-085-199B-11 Sequence 11, Appl
44	131	7.1	1093	5	PCT-US93-03077-1 Sequence 1, Appl
45	130.5	7.1	396	1	US-08-430-024-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-235-836C-110  
Sequence 110, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Botrellia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 52.3%; Score 964; DB 4; Length 466;  
Best local Similarity 59.6%; Pred. No. 7e-62;  
Matches 229; Conservative 31; Mismatches 92; Indels 32; Gaps 9;  
QY 1 MACNNSGKCDNTSANSADSEVKPNLJEISKRTTDSNAVLAVKEVALLSIDEIAAKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDTSNVLAVKEVEALLSIDETIAKA 76  
Qy 61 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120  
Db 77 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136  
Qy 121 KKEKHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 180  
Db 137 KKAHHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 196  
Qy 181 ELTSPVAESP--AMVNSGKDGNTSANSADSVKGNPLTEISKRTIESNA--VVLAVKE 236  
Db 197 ELTSPVAESP--AMVNSGKDGNTSANSADSVKGNPLTEISKRTIESNA--VVLAVKE 253  
Qy 237 VETLTS-----IDELAKAIGKIKNDVSLDNEADHNSLISGAYLT--SMLTKKIS 287  
Db 254 LELKGTSDKNNKSGVLEGVKADKSKVLTIS-----DILQOTLLEVKEDGKTLVSKRYT 308  
Qy 288 AKKSGELKAEIEKAKKCEEFITAKLGEHTDLGKGVTDNNAKKAITKTNNKTGADG 347  
Db 309 S-KKSSSTEKFEKNGEVSEKIIIRADG--TRLEYTGIKSDGSKAKKAVLKGIVLEGTLT 365  
Qy 348 LEKLFESY-----KNLSKA 363  
Db 366 AEKTLTVKEGCTVTLKNNKSGE 389

RESULT 2  
US-08-235-836C-30  
; Sequence 30, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
; TITLE OF INVENTION: Borrelia Polypeptides and Uses therefor  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235,836C  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 210 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-836C-30

Query Match 50.6%; Score 932; DB 4; Length 210;  
Best Local Similarity 98.4%; Pred. No. 4.9e-60;  
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRTIDTSNVLAVKEVEALLSIDETIAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDTSNVLAVKEVEALLSIDETIAKA 76  
Qy 61 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120  
Db 77 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136  
Qy 121 KKEKHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 180  
Db 137 KKEKHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 196  
Qy 181 ELTSPVAESP 191  
Db 197 ELTSPVAESP 207

RESULT 3  
US-09-196-293-15  
; Sequence 15, Application US/09196293  
; Patent No. 6183755  
; GENERAL INFORMATION:  
; APPLICANT: Fuchs, Renate  
; APPLICANT: Motz, Manfred  
; APPLICANT: Sautscheck, Erwin  
; APPLICANT: Wilsch, Bettina  
; APPLICANT: Preac-Mursic, Vera  
; TITLE OF INVENTION: Active proteins from Borrelia  
; FILE REFERENCE: 738.001052  
; CURRENT APPLICATION NUMBER: US/09/196,293  
; EARLIER FILING DATE: 1998-11-19  
; EARLIER APPLICATION NUMBER: US 08/209,603  
; EARLIER FILING DATE: 1994-03-10  
; EARLIER APPLICATION NUMBER: US 07/862,535  
; EARLIER FILING DATE: 1992-06-19  
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
; EARLIER FILING DATE: 1990-12-21  
; EARLIER APPLICATION NUMBER: DE P39 42 728.5  
; EARLIER FILING DATE: 1989-12-22  
; EARLIER APPLICATION NUMBER: DE P40 18 988.0  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; TYPE: PRP  
; LENGTH: 209  
; ORGANISM: Borrelia burgdorferi  
; US-09-196-293-15

Query Match 50.4%; Score 928; DB 4; Length 209;  
Best Local Similarity 98.4%; Pred. No. 9.5e-60;  
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRTIDTSNVLAVKEVEALLSIDETIAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDTSNVLAVKEVEALLSIDETIAKA 76  
Qy 61 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120  
Db 77 ICKKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136  
Qy 121 KKEKHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 180  
Db 137 KKEKHTDLGKGVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSYK 196  
Qy 181 ELTSPVAESP 191  
Db 197 ELTSPVAESP 207

Db 197 ELTSPVVAESP 207

## RESULT 4

US-08-158-353-3  
; Sequence 3, Application US/08158353  
; Patent No. 5620862

## GENERAL INFORMATION:

APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-9540  
TELEFAX: 617-861-6240

## INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-3

## Query Match

Best Local Similarity 50.4%; Score 928; DB 1; Length 210;  
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 MACNNGKCGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKA 60

17 ISCNNGKCGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKA 76

61 IGRKIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFN 120

77 IGRKIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFN 136

121 KLEKHTDGLKEGVTDADAKKALILKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

137 KLEKHTDGLKEGVTDADAKKALILKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

QY 181 ELTSPVVAESP 191

Db 197 ELTSPVVAESP 207

## RESULT 5

US-08-235-836C-122  
; Sequence 122, Application US/08235836C  
; Patent No. 6248362

## GENERAL INFORMATION:

APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729

## INFORMATION FOR SEQ ID NO: 122:

SEQUENCE CHARACTERISTICS:  
LENGTH: 588 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-122

## Query Match

Best Local Similarity 50.3%; Score 927; DB 4; Length 588;  
Matches 193; Conservative 4; Mismatches 10; Indels 10; Gaps 2;

QY 4 NNSGKCGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKAIGK 63

Db 286 NNSGKCGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKAIGK 345

QY 64 KIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFNK 123

Db 346 KIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFNK 405

QY 124 EKHTDGLKEGVTDADAKKALILKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 183

Db 406 EKHTDGLKEGVTDADAKKALILKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 465

QY 184 SPVVAESP-----AMVNN-----SGKDGNTSANSAD 210

Db 466 SPVVAESP-----AMVNN-----SGKDGNTSANSAD 502

## RESULT 6

US-08-209-603E-15  
; Sequence 15, Application US/08209603E  
; Patent No. 6248338

## GENERAL INFORMATION:

APPLICANT: FUCHS, RENATE  
APPLICANT: WILKE, BETTINA  
APPLICANT: PREAC-MURSIC, VERA  
APPLICANT: MORTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210  
TYPE: AMINO ACID  
MOLECULE TYPE: LINEAR  
DESCRIPTION: PROTEIN  
HYPOTHETICAL: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: B. BURGDORFERI  
IMMEDIATE SOURCE:  
LIBRARY: DSM 5662  
POSITION IN GENOME: N/A  
FEATURE:  
IDENTIFICATION METHOD: amino acid analysis  
PUBLICATION INFORMATION: N/A  
US-08-209-603E-15

Query Match  
Best Local Similarity 50.0%; Score 921; DB 4; Length 210;  
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVALLSSIDEIAKA 60  
DB :|||||  
QY 61 IGKIHONGLDTEYNNHNSLLAGAYASTLIKOKLDGLKNEGKLEKIDAAKCCSEFTN 76  
DB :|||||  
QY 77 IGKIHONGLDTEYNNHNSLLAGAYASTLIKOKLDGLKNEGKLEKIDAAKCCSEFTN 120  
DB :|||||  
QY 121 KLEKHTDGLKGVYDADAKKAILKTNGTKKGAEEGLKFESVEVLSKAKEMLANSVK 180  
DB :|||||  
QY 137 KLEKHTDGLKGVYDADAKKAILKTNGTKKGAEEGLKFESVEVLSKAKEMLANSVK 196  
QY 181 ELTSPVAESP 191  
DB :|||||  
QY 197 ELTSPVAESP 207

RESULT 7  
US-08-235-836C-107  
Sequence 107, Application US/08235836C  
Patent No. 6248562

## GENERAL INFORMATION:

APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
NUMBER OF INVENTION: Borrella Polypeptides and Uses Therefor  
CORRESPONDENCE ADDRESS: 144  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
US-08-235-836C-107

Query Match  
Best Local Similarity 49.7%; Score 915; DB 4; Length 466;  
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVALLSSIDEIAKA 63  
DB :|||||  
QY 64 KIHONGLDTEYNNHNSLLAGAYASTLIKOKLDGLKNEGKLEKIDAAKCCSEFTN 123  
DB :|||||  
QY 124 EKHTDGLKGVYDADAKKAILKTNGTKKGAEEGLKFESVEVLSKAKEMLANSVK 183  
DB :|||||  
QY 396 AKHTDGLKGVYDADAKKAILKTNGTKKGAEEGLKFESVEVLSKAKEMLANSVK 455  
QY 184 SPYVAESP 191  
DB :|||||  
QY 456 SPYVAESP 463

RESULT 8  
US-08-158-353-2  
Sequence 2, Application US/08158353  
Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF INVENTION: Disease  
CORRESPONDENCE ADDRESS: 7  
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.



Query Match 33.2%; Score 612; DB 4; Length 212;

Best Local Similarity 70.5%; Pred. NO. 4,8e-37;  
Matches 136; Conservative 12; Mismatches 43; Indels 2; Gaps 2

Oy	60	AIGKKIHQNNGLDTEYHNGSLLAGAYASTLKIKOKIDGKN-EGKLEKIDAARKCSETF	118
Dd	77	AIGKIDONNGLAALNNONGSLLAGAVASLITLERLSKLNEBELKTEIATAKCKSEEF	136
Oy	119	TNKLRKEKHTDGRKGCVDDADAKELTKNGTKRKGAEFLKREBSVSVLSKAAREMIANS	178
Dd	137	TNRKLSGGADGSKODADDNDHAKAIILTKHTDTDGAREFEDVESVGILLKAAOVALTNS	186
Oy	179	VKELTSPVVAESP	191
Dd	197	VKELTSPVVAESP	209

RESULT 13  
08-209-

Sequence 11, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION

GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILSKE, BETTINA  
APPLICANT: PREAC-MURISIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHEK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELIA BURDORFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:

1 ADDRESSSEE: BROOKS HAIDT HAFNER & DE  
2 STREET: 99 PARK AVENUE  
3 CITY: NEW YORK  
4 STATE: NY  
5 COUNTRY: USA  
6 ZIP: 10016  
7  
8 COMPUTER READABLE FORM:  
9 MEDIUM TYPE: 3.5" FLOPPY DISC  
10 COMPUTER: AT&T - IBM COMPATIBLE  
11 OPERATING SYSTEM: MS-DOS Version 6.2.2

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-P-P-  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (412) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEO ID NO: 11:  
SEQUENCE CHARACTERISTICS:

```

: IMMEDIATE SOURCE:
: LIBRARY: DSM 5662
: POSITION IN GENOME: N/A
: FEATURE:
: IDENTIFICATION METHOD: amino acid analysis
: PUBLICATION INFORMATION: N/A
: ANTI-SENSE: N/A
: FRAGMENT TYPE: INTERNAL
: ORIGINAL SOURCE:
US-08-209-603E-11

```

Query Match	33.2%	Score 612	DB 4	Length 212
Best Local Similarity	70.5%	Pred. No.	4.8e-37	
Matches 136	Conservative 12	Mismatches 43	Indels 2	Gaps 2

[illegible]

RESULT 14  
TIS-08-235-

; Sequence 34, Application US/08235836C  
; Patent No. 6248562

GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BUL93-288  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-34

Query Match 33.1%; Score 609; DB 4; Length 212;  
Best Local Similarity 69.4%; Pred. No. 7.9e-37;  
Matches 134; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSADESVPKPNLTETSKITDSNAVLLAVKEVEALLSSIDEIAR 59  
Db 17 ISCSNSKGGDASTNPDESAGPNLTETSKITDSNAFLAVKEVEALLSSIDEIAR 76  
QY 60 AIGKIHQNNGLDTEYVHNGSLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 118  
Db 77 AIGKIDNNNGLAALNNQNSLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 136  
119 TNKIKKHTDLGKGVTDADAKKAIILKTNGTKTGAEEGLKFESEVLSKAKEMLANS 178  
137 TNKIKSGHADLGQDADDDHAKAAILKTHTATDGAKEFDFEVEGLKKAQVALTNS 196  
QY 179 VKELTSPVAESP 191  
Db 197 VKELTSPVAESP 209

RESULT 15

US-08-235-836C-32  
Sequence 32, Application US/08235836C  
Patent No. 6248562

GENERAL INFORMATION:

APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/235, 836C

FILING DATE: 29-APR-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191

FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 209 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-32

Matches 131; Conservative 19; Mismatches 39; Indels 3; Gaps 2;

QY 1 MACNSGKDGNTSANSADESVPKPNLTETSKITDSNAVLLAVKEVEALLSSIDEIAR 60  
Db 17 ISCSNSG--GDTASTNPDESAGPNLTETSKITDSNAFLAVKEVEALLSSIDEIAR 74  
QY 61 ICKTIHQNNGLDTEYVHNGSLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 119  
Db 75 ICKVIHQNNGLNANMGQNSLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 134  
QY 120 NKLKREKHTDLGKGVTDADAKKAIILKTNGTKTGAEEGLKFESEVLSKAKEMLANS 179  
Db 135 NKLKSHQGLGVAATAITDHAKEAILKSNPTKDGAKALKDSESEVLSKAAQALANSV 194  
QY 180 KELTSPVAESP 191  
Db 195 KELTSPVAESP 206

Search completed: March 18, 2002, 09:55:32  
Job time: 345 sec

Query Match 32.6%; Score 600.5; DB 4; Length 209;  
Best Local Similarity 68.2%; Pred. No. 3.2e-36;





**THIS PAGE BLANK**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:32 ; Search time 621.2 Seconds

(without alignments)  
167.613 Million cell updates/sec

Title: US-09-596-746A-24

Sequence: 1 MACNNSGKDGNTSANSADSE.....KNLSKAKEMLTNSVKELTS 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 segs, 277657034 residues

Database : Pending Patents, AA, Main:\*

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1842	100.0	375	19	US-09-596-746A-24
2	1837	99.7	374	19	US-09-596-746-24
3	1834	99.6	398	19	US-09-596-746-60
4	1834	99.6	399	19	US-09-596-746A-60
5	1819.5	98.8	401	19	US-09-596-746-80
6	1819.5	98.8	402	19	US-09-596-746A-80
7	1810.5	98.3	378	19	US-09-596-746-28
8	1810.5	98.3	378	19	US-09-596-746A-28
9	1597.5	86.7	374	19	US-09-596-746A-26

10	1592.5	86.5	373	19	US-09-596-746-26	Sequence 26, Appl
11	1589.5	86.3	397	19	US-09-596-746-62	Sequence 62, Appl
12	1589.5	86.3	398	19	US-09-596-746A-62	Sequence 82, Appl
13	1575	85.5	400	19	US-09-596-746-82	Sequence 82, Appl
14	1575	85.5	401	19	US-09-596-746A-82	Sequence 82, Appl
15	1566	85.0	377	19	US-09-596-746-30	Sequence 30, Appl
16	1566	85.0	377	19	US-09-596-746A-30	Sequence 30, Appl
17	1534	83.3	371	19	US-09-596-746A-34	Sequence 34, Appl
18	1529	83.0	370	19	US-09-596-746-34	Sequence 34, Appl
19	1526	82.8	394	19	US-09-596-746-64	Sequence 64, Appl
20	1526	82.8	395	19	US-09-596-746A-64	Sequence 64, Appl
21	1498	81.3	369	19	US-09-596-746A-40	Sequence 40, Appl
22	1493	81.1	368	19	US-09-596-746-40	Sequence 40, Appl
23	1493	81.1	385	19	US-09-596-746A-56	Sequence 56, Appl
24	1492	81.0	560	19	US-09-596-746-52	Sequence 52, Appl
25	1492	81.0	560	19	US-09-596-746A-52	Sequence 52, Appl
26	1490	80.9	392	19	US-09-596-746-72	Sequence 72, Appl
27	1490	80.9	393	19	US-09-596-746A-72	Sequence 72, Appl
28	1488	80.8	384	19	US-09-596-746-56	Sequence 56, Appl
29	1485	80.6	408	19	US-09-596-746-78	Sequence 78, Appl
30	1485	80.6	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289.5	70.0	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284.5	69.7	369	19	US-09-596-746-36	Sequence 36, Appl
33	1281.5	69.6	393	19	US-09-596-746-66	Sequence 66, Appl
34	1281.5	69.6	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1253.5	68.1	368	19	US-09-596-746A-42	Sequence 42, Appl
36	1248.5	67.8	367	19	US-09-596-746-42	Sequence 42, Appl
37	1245.5	67.6	391	19	US-09-596-746-74	Sequence 74, Appl
38	1245.5	67.6	392	19	US-09-596-746A-74	Sequence 74, Appl
39	1217.5	66.1	370	19	US-09-596-746A-32	Sequence 32, Appl
40	1213.5	65.9	378	19	US-09-596-746-54	Sequence 54, Appl
41	1213.5	65.9	378	19	US-09-596-746A-54	Sequence 54, Appl
42	1212.5	65.8	369	19	US-09-596-746-32	Sequence 32, Appl
43	1209.5	65.7	393	19	US-09-596-746-68	Sequence 68, Appl
44	1209.5	65.7	394	19	US-09-596-746A-68	Sequence 68, Appl
45	1205.5	65.4	401	19	US-09-596-746-70	Sequence 70, Appl

#### ALIGNMENTS

RESULT 1  
US-09-596-746A-24  
; Sequence 24, Application US/09596746A  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seinstost, Gerald  
; APPLICANT: Dykhuisen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease In Humans  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746A  
; CURRENT FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 375  
; TYPE: PRN  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746A-24

Query Match 100.0%; Score 1842; DB 19; Length 375;  
Best Local Similarity 100.0%; Pred. No. 8; 8e-125;  
Matches 375; Conservative 0; Mismatches 0; Gaps 0;  
QY 1 MACNNSGKDGNTSANSADSEVKGNLTREISKITDSNAVLVAKEVALLSIDEIAKA 60

```

Db 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 60
QY 61 IGKIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 120
Db 61 IGKIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 120
QY 121 KKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 180
Db 121 KKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 240
Db 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 240
QY 241 LTSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIE 300
Db 241 LTSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIE 300
QY 301 KAKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSK 360
Db 301 KAKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSK 360
QY 361 AKEMLTNSVKELTS 375
Db 361 AKEMLTNSVKELTS 375

RESULT 2
US-09-596-746-24
; Sequence 24, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OspC Chimera
US-09-596-746-24

```

```

Query Match 99.7%; Score 1837; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 2e-124;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 61
Db 1 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 61
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 120
Db 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 120
QY 122 LKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 181
Db 122 LKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 181
QY 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 241
Db 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 241

```

```

Db 181 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 240
QY 242 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIEK 301
Db 241 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIEK 301
QY 302 AKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSKA 361
Db 301 AKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSKA 361
QY 362 AKEMLTNSVKELTS 375
Db 361 AKEMLTNSVKELTS 374

RESULT 3
US-09-596-746-60
; Sequence 60, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-60

```

```

Query Match 99.6%; Score 1834; DB 19; Length 398;
Best Local Similarity 99.7%; Pred. No. 3.6e-124;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKAI 84
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 121
Db 85 GKRIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKEKIDAKKCSFTFN 144
QY 122 LKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 181
Db 145 LKEKHTDLGKEGYTDADAKAIIKTNGTKGAEEGLFESVEVLSAAKEMLANSVKE 181
QY 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 241
Db 205 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVVLAVKEVETL 264
QY 242 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIEK 301
Db 265 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAVLLSNLTTKISAINDSGELKAEIEK 324
QY 302 AKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSKA 361
Db 325 AKKCEEFPAKLKGHEHTDLGKEGYTDNNAKAKIILKTNNDKTKGADELEKLFESVKNLSKA 361
QY 362 AKEMLTNSVKELTS 375
Db 385 AKEMLTNSVKELTS 398

```

```
RESULT 4
US-09-596-746a-60
; Sequence 60, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuitzen, Dantai
; APPLICANT: Luf, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospc Chimera
09-596-746a-60

Query Match          99.6%; Score 1834; DB 19; Length 399;
Best Local Similarity 99.7%; Pred. No. 3,6e-123;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKTTSDSNAVLLAVEBALLSSIDEIAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKTTSDSNAVLLAVEBALLSSIDEIAKAI 85
QY 62 GKRIHONNGLDPEYHNHNSLAGAVAI STLTKQKLDGLKNEGKLEKIDAKKCEFTFNK 121
DB 86 GKRIHONNGLDPEYHNHNSLAGAVAI STLTKQKLDGLKNEGKLEKIDAKKCEFTFNK 145
QY 122 LKEKHTDLGKEGVTDDAKAEALIKTNGTKGAEEGLKFESEVLSKAKEMLANSVKE 181
DB 146 LKEKHTDLGKEGVTDDAKAEALIKTNGTKGAEEGLKFESEVLSKAKEMLANSVKE 205
QY 182 LTPVVAESPRAWNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLLAVEETLL 241
DB 206 LTPVVAESPRAWNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLLAVEETLL 265
QY 242 TSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLITRKISAIKDSGELKAEIEK 301
DB 266 TSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLITRKISAIKDSGELKAEIEK 325
QY 302 AKKCESEFTAKLGEHTDLGKEGVTDDNAKKAIIKTNDKTKGADELEKLFESVKNLSKA 361
DB 326 AKKCESEFTAKLGEHTDLGKEGVTDDNAKKAIIKTNDKTKGADELEKLFESVKNLSKA 385
QY 362 AKEMLTNSVKEIETS 375
DB 386 AKEMLTNSVKEIETS 399

RESULT 5
US-09-596-746-80
; Sequence 80, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuitzen, Dantai
; APPLICANT: Luf, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-80

Query Match          98.8%; Score 1819.5; DB 19; Length 401;
Best Local Similarity 98.7%; Pred. No. 4,1e-123;
Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKTTSDSNAVLLAVEBALLSSIDEIAKAI 61
DB 25 SCNNSGKDGNTSANSADSVKGPNTLEISKTTSDSNAVLLAVEBALLSSIDEIAKAI 84
QY 62 GKRIHONNGLDPEYHNHNSLAGAVAI STLTKQKLDGLKNEGKLEKIDAKKCEFTFNK 121
DB 85 GKRIHONNGLDPEYHNHNSLAGAVAI STLTKQKLDGLKNEGKLEKIDAKKCEFTFNK 144
QY 122 LKEKHTDLGKEGVTDDAKAEALIKTNGTKGAEEGLKFESEVLSKAKEMLANSVKE 181
DB 145 LKEKHTDLGKEGVTDDAKAEALIKTNGTKGAEEGLKFESEVLSKAKEMLANSVKE 204
QY 182 LTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLLAVEVE 238
DB 205 LTPVVAESPKRPMSVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLLAVEVE 264
QY 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLITRKISAIKDSGELKAE 298
DB 265 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLITRKISAIKDSGELKAE 324
QY 299 IEKAKKCESEFTAKLGEHTDLGKEGVTDDNAKKAIIKTNDKTKGADELEKLFESVKNL 358
DB 325 IEKAKKCESEFTAKLGEHTDLGKEGVTDDNAKKAIIKTNDKTKGADELEKLFESVKNL 384
QY 359 SKAKEMLTNSVKEIETS 375
DB 385 SKAKEMLTNSVKEIETS 401

RESULT 6
US-09-596-746a-80
; Sequence 80, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuitzen, Dantai
; APPLICANT: Luf, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 402
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746a-80

Query Match          98.8%; Score 1819.5; DB 19; Length 402;
Best Local Similarity 98.7%; Pred. No. 4,1e-123;
Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;
```

Db 26 SCNNSGDGTNSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVEALLSSIDEIAKAI 85  
QY 62 GKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 121  
Db 86 GKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 145  
QY 122 LKEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 181  
Db 146 LKEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 205  
QY 182 LTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 238  
Db 206 LTSPVVAESPKPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 265  
QY 239 TLTSIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 298  
Db 266 TLTSIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 325  
QY 299 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKNL 358  
Db 326 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKNL 385  
QY 359 SKAKEMLTNSVKELTS 375  
Db 386 SKAKEMLTNSVKELTS 402

RESULT 7  
US-09-596-746-28  
; Sequence 28, Application US/09596746  
; GENERAL INFORMATION:  
; APPLICANT: Datwyler, Raymond J.  
; APPLICANT: Selmost, Gerald  
; APPLICANT: Dykhuzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT FILING DATE: US/09/596/746  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; NO ID NO 28  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746-28

Query Match 98.3%; Score 1810.5; DB 19; Length 378;  
Best Local Similarity 98.1%; Pred. No. 1.7e-122;  
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVEALLSSIDEIAKA 60  
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVEALLSSIDEIAKA 60  
QY 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 120  
Db 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 120  
QY 121 KLEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180  
Db 121 KLEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180  
QY 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 237  
Db 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 237

Db 181 ELTSPVVAESPKPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 240  
QY 238 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 297  
Db 241 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 300  
QY 298 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKN 357  
Db 301 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKN 360  
QY 358 LSKAKEMLTNSVKELTS 375  
Db 361 LSKAKEMLTNSVKELTS 378

RESULT 8  
US-09-596-746A-28  
; Sequence 28, Application US/09596746A  
; GENERAL INFORMATION:  
; APPLICANT: Datwyler, Raymond J.  
; APPLICANT: Selmost, Gerald  
; APPLICANT: Dykhuzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT FILING DATE: US/09/596/746A  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746A-28

Query Match 98.3%; Score 1810.5; DB 19; Length 378;  
Best Local Similarity 98.1%; Pred. No. 1.7e-122;  
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVEALLSSIDEIAKA 60  
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVEALLSSIDEIAKA 60  
QY 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 120  
Db 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCCSEFTTN 120  
QY 121 KLEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180  
Db 121 KLEKHTDGLGEGYTDADAKAAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180  
QY 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 237  
Db 181 ELTSPVVAESPKPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVAVLLAVEVE 240  
QY 238 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 297  
Db 241 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAYILSNLITTKISAKISGELKAE 300  
QY 298 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKN 357  
Db 301 IEKAKCSEFTTAKLKGHTDGLGEGYTDNNAKAILKTNDKTKGADELKLFESVKN 360  
QY 358 LSKAKEMLTNSVKELTS 375  
Db 361 LSKAKEMLTNSVKELTS 378

```

RESULT 9
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-26

Query Match      86.7%; Score 1597.5; DB 19; Length 374;
Best Local Similarity 87.7%; Pred. No. 4.3e-107;
Matches 328; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVALSSIDEIAAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVALSSIDEIAAKA 60
QY 61 ICKKIHQNNGLDTEYNNHNSLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCEFTFN 120
DB 61 ICKKIHQNNGLDTEYNNHNSLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCEFTFN 120
QY 121 KKEKHTDGLKGEVTDADAKKALTKNGTKGAEBELGFESVEVLSKAKEMLANSVK 180
DB 121 KKEKHTDGLKGEVTDADAKKALTKNGTKGAEBELGFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEVETL 240
DB 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEVETL 240
QY 241 LVSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKRIISAIDSGELKAEI 299
DB 241 LVSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKRIISAIDSGELKAEI 299
QY 300 EKAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 359
DB 300 EKAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 359
QY 361 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360
DB 361 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360
QY 360 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360
DB 360 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360
QY 361 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360
DB 361 KAKKCEFTKLEGEHQGLGIEVTDENAKKALITDAKKGAAELKLEKRAVENLA 360

RESULT 10
US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-26

Query Match      86.7%; Score 1597.5; DB 19; Length 374;
Best Local Similarity 87.7%; Pred. No. 4.3e-107;
Matches 328; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```

```

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-26

Query Match      86.5%; Score 1592.5; DB 19; Length 373;
Best Local Similarity 87.7%; Pred. No. 9.9e-107;
Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVALSSIDEIAAKA 61
DB 1 ACNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVALSSIDEIAAKA 60
QY 62 GKRIHQNNGLDTEYNNHNSLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCEFTFN 121
DB 62 GKRIHQNNGLDTEYNNHNSLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCEFTFN 120
QY 122 LKKEKHTDGLKGEVTDADAKKALTKNGTKGAEBELGFESVEVLSKAKEMLANSVK 181
DB 122 LKKEKHTDGLKGEVTDADAKKALTKNGTKGAEBELGFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEVETL 241
DB 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEVETL 240
QY 242 TSDIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKRIISAIDSGELKAEI 300
DB 242 TSDIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKRIISAIDSGELKAEI 300
QY 301 KAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 360
DB 301 KAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 360
QY 361 KAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 360
DB 361 KAKKCEFTKAKLKEGHTDGLKGEVTDADAKKALTKNNDRKKGADLEKLEESYKNS 360

RESULT 11
US-09-596-746-62
; Sequence 62, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRP
; ORGANISM: OspC Chimera
US-09-596-746-62

```





Db 385 LAKAKEMLANSVKEL 400

## RESULT 14

US-09-596-746A-82  
Sequence 82, Application US/09596746A

## GENERAL INFORMATION:

APPLICANT: Dettwyler, Raymond J.  
APPLICANT: Selmozt, Gerald  
APPLICANT: Dykhuisen, Daniel  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746A  
CURRENT FILING DATE: 2000-06-19  
PRIOR APPLICATION NUMBER: US 60/140,042  
PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 82

LENGTH: 401  
TYPE: PRT  
ORGANISM: ospC Chimera  
US-09-596-746A-82

Query Match 85.5%; Score 1575; DB 19; Length 401;  
Best Local Similarity 86.4%; Pred. No. 2e-105;  
Matches 335; Conservative 22; Mismatches 25; Indels 4; Gaps 2;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKAI 61  
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKAI 85  
QY 62 GKRIHONGLDPEYNNHNSGLAGAVAI STLKOKLKGKLEKIDAAKCSFTFNK 121  
DB 86 GKRIHONGLDPEYNNHNSGLAGAVAI STLKOKLKGKLEKIDAAKCSFTFNK 145  
QY 122 LKREKHTDLGEGVTDADAKFAILKTNGTKTGABEILGFESVEVLSKAEMLANSVKE 181  
DB 146 LKREKHTDLGEGVTDADAKFAILKTNGTKTGABEILGFESVEVLSKAEMLANSVKE 205  
QY 182 LTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 238  
DB 206 LTPVVAESPKKPSMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 265  
QY 239 TLATSTIDELA-KAIGKKIKNDVSLDNEADHNSGLISGAVIISLITKIKSAIDSGELKA 297  
DB 266 TLATSTIDELATKAIGKKIQONGSLAVEAGHNGTLAGAVIISLITKIKLGLKSEKLE 325  
QY 298 EIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 357  
DB 326 KIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 385  
QY 358 LSKAAKEMLTNSVKEL 373  
DB 386 LAKAKEMLANSVKEL 401

## RESULT 15

US-09-596-746-30  
Sequence 30, Application US/09596746

## GENERAL INFORMATION:

APPLICANT: Dettwyler, Raymond J.  
APPLICANT: Selmozt, Gerald  
APPLICANT: Dykhuisen, Daniel  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
TITLE OF INVENTION: Borrelia afzelii that Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 30  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: OspC Chimera  
US-09-596-746-30

Query Match 85.0%; Score 1566; DB 19; Length 377;  
Best Local Similarity 85.9%; Pred. No. 8.3e-105;  
Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKA 60  
DB 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKA 60  
QY 61 IGKRIHONGLDPEYNNHNSGLAGAVAI STLKOKLKGKLEKIDAAKCSFTFNK 120  
DB 61 IGKRIHONGLDPEYNNHNSGLAGAVAI STLKOKLKGKLEKIDAAKCSFTFNK 120  
QY 121 LKREKHTDLGEGVTDADAKFAILKTNGTKTGABEILGFESVEVLSKAEMLANSVK 180  
DB 121 LKREKHTDLGEGVTDADAKFAILKTNGTKTGABEILGFESVEVLSKAEMLANSVK 180  
QY 181 ELTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 237  
DB 181 ELTPVVAESPKKPSMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 240  
QY 238 ETLTSTIDELA-KAIGKKIKNDVSLDNEADHNSGLISGAVIISLITKIKSAIDSGELK 296  
DB 241 ETLTSTIDELATKAIGKKIQONGSLAVEAGHNGTLAGAVIISLITKIKLGLKSEKLE 300  
QY 297 AEIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 356  
DB 301 EIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 360  
QY 357 NLKAAKEMLTNSVKEL 373  
DB 361 NLKAAKEMLTNSVKEL 377

Search completed: March 18, 2002, 10:08:42  
Job time: 970 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:02 ; Search time 81.02 Seconds  
(without alignments)  
295,677 Million cell updates/sec

Title: US-09-596-746A-24

Perfect score: 1842  
Sequence: 1 MACNNSCKDGTSTANSADSE.....KNLSKAKEMLTNSVKELTS 375

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US11\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US12\_NEW\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US13\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	674.5	36.6	211	US-09-974-992-5	Sequence 5, Appl
2	614	33.3	212	US-09-974-992-7	Sequence 7, Appl
3	587.5	31.9	207	US-09-974-992-3	Sequence 3, Appl
4	175	9.5	1014	US-09-708-427-19883	Sequence 19883, A
5	175	9.5	1018	US-09-708-427-19882	Sequence 19882, A
6	175	9.5	1269	US-09-708-427-19881	Sequence 19881, A
7	169.5	9.2	1313	US-09-708-427-15044	Sequence 15044, A
8	169	9.2	1881	US-10-032-851-7646	Sequence 7646, Ap
9	169	9.2	1881	US-10-072-851-15590	Sequence 15590, A
10	168.5	9.1	1504	US-09-708-427-15045	Sequence 15045, A
11	158	8.6	589	US-09-748-875-14	Sequence 15045, A
12	158	8.6	690	US-09-748-875-61	Sequence 61, Appl
13	158	8.6	691	US-09-748-875-1	Sequence 61, Appl
14	158	8.6	701	US-09-748-875-62	Sequence 62, Appl
15	158	8.6	707	US-09-748-875-2	Sequence 2, Appl
16	158	8.6	837	US-09-815-242-5883	Sequence 5883, Ap
17	158	8.6	837	US-10-072-851-13080	Sequence 5883, Ap
18	158	8.6	875	US-09-815-242-13080	Sequence 13080, A
19	158	8.6	925	US-10-072-851-13080	Sequence 13080, A
20	158	8.6	929	US-09-748-875-60	Sequence 60, Appl
21	157.5	8.6	1725	US-10-037-182-12	Sequence 12, Appl
22	157.5	8.6	1786	US-09-938-275-7	Sequence 7, Appl
23	157.5	8.6	1786	US-10-037-182-10	Sequence 10, Appl
24	154.5	8.4	1144	US-09-708-427-15046	Sequence 15046, A
25	154.5	8.4	2434	US-09-815-242-5835	Sequence 5835, Ap

26	154.5	8.4	2434	US-10-072-851-5835	Sequence 5835, Ap
27	154.5	8.4	6281	US-09-815-242-12996	Sequence 12996, A
28	154.5	8.4	6281	US-10-072-851-12996	Sequence 12996, A
29	152.5	8.3	1703	US-09-708-427-21161	Sequence 21161, A
30	152.5	8.3	1710	US-09-708-427-21161	Sequence 21160, A
31	152.5	8.3	1744	US-09-708-427-21159	Sequence 21159, A
32	152.5	8.3	2437	US-09-815-242-5834	Sequence 5834, Ap
33	152.5	8.3	2437	US-10-072-851-5834	Sequence 5834, Ap
34	150.5	8.2	670	US-09-748-875-63	Sequence 63, Appl
35	150.5	8.2	1690	US-09-614-150-10224	Sequence 10224, A
36	150.5	8.2	1690	US-09-614-150-10311	Sequence 10311, A
37	150	8.1	805	US-09-708-427-3730	Sequence 3730, Ap
38	150	8.1	841	US-09-708-427-3729	Sequence 3729, Ap
39	146	7.9	711	US-09-748-875-3	Sequence 3, Appl
40	146	7.9	1765	US-10-037-182-8	Sequence 8, Appl
41	146	7.9	1786	US-09-938-275-6	Sequence 6, Appl
42	146	7.9	1786	US-09-873-676-113	Sequence 113, App
43	146	7.9	1786	US-09-561-7098-9	Sequence 9, Appl
44	146	7.9	1786	US-10-037-182-6	Sequence 6, Appl
45	146	7.9	2086	US-09-815-242-5639	Sequence 5639, Ap

## ALIGNMENTS

RESULT 1  
US-09-974-992-5  
; Sequence 5, Application US/0974992  
; GENERAL INFORMATION:  
; APPLICANT: Mathiesen, Marianne J.  
; APPLICANT: Theisen, Michael  
; APPLICANT: Holm, Arne  
; TITLE OF INVENTION: Novel Opc-derived peptide fragments  
; FILE REFERENCE: 459-666p  
; CURRENT APPLICATION NUMBER: US/09/974, 992  
; CURRENT FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: 09/180, 089  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: PCT/DK97/00203  
; PRIOR FILING DATE: 1997-05-02  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-974-992-5

Query Match 36.6%; Score 674.5; DB 6; Length 211;  
Best Local Similarity 76.7%; Pred. No. 2.4e-39;  
Matches 148; Conservative 11; Mismatches 31; Indels 3; Gaps 3;  
QY 1 MACNNSCKDGTSTANSADSEVKGPNLTSEIKKTTDSNAVLLAVEFALLSSIDEIARA 60  
DB 17 ISCNNSCKDGTSTANSADSEVKGPNLTSEIKKTTDSNAVLLAVEFALLSSIDEIARA 75  
QY 61 IGKRIHONNGIDLEYHNHNSILAGAYVISTLIKOKIDGLK-NBGLKKEIDAKKCSFTT 119  
DB 76 IGKRIKNDGSLGDEANNESELGAYVISTLIKOKIDGLK-NBGLKKEIDAKKCSFTT 135  
QY 120 NKLEKHTDYGKGVYDADKAEALIKNGT-KTKGAELCKLPESVYLSKAKKEMLANS 178  
DB 136 TKLKDNHQAQIGGVYDADKAEALIKNGT-KTKGAELCKLPESVYLSKAKKEMLANS 195  
QY 179 VKELTSPVVAESP 191  
DB 196 VKELTSPVVAESP 208  
RESULT 2  
US-09-974-992-7

```

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSCP-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrella afzei11
US-09-974-992-7

```

```

Query Match          33.3%; Score 614; DB 6; Length 212;
Best Local Similarity 69.9%; Pred. No. 3,4e-35;
Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNT--SANSADSVKGNLTETSKRTTDSNAVLAVKEVALSSIDEIAK 59
   :||||| : : : ||| ||||| ||||| ||||| ||||| : ||||| : ||||| :
Db 17 ISCNNSGKGSASTNPADSEAKGNPLTETSKRTTDSNAVLAVKEVALSSIDEIAK 76
   :||||| : : : ||| ||||| ||||| ||||| ||||| : ||||| : ||||| :
QY 60 AIGKIHONNGLDTEYHNHNSLAGAYAISTLIKQKIDGLKN--EGLKEKIDAAKCSSEF 118
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 77 AIGKIDNNNGSLALNNQNSLAGAYAISTLIKTEKLSKKNLEELKTEIAKAKKCSSE 136
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 119 TNKLEKHTDLGKEGVTDADAKFAILKTNGTKGAEELGKFESEVYLSKAKKEMLAN 178
   |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| :
Db 137 TNKLSGHADLGQADTDHAKAAILKTATTTDKGAKERDLESEVGLKAQAVALTNS 196
   |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| :
QY 179 VKELTSPVVAESP 191
   |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| :
Db 197 VKELTSPVVAESP 209
   |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| :

```

```

RESULT 3
US-09-974-992-3
Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSCP-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrella garinii
US-09-974-992-3

```

```

Query Match          31.9%; Score 587.5; DB 6; Length 207;
Best Local Similarity 66.5%; Pred. No. 2,2e-33;
Matches 127; Conservative 24; Mismatches 37; Indels 3; Gaps 2;

```

```

QY 1 MACNNSGKDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKEVALSSIDEIAKA 60
   :||||| : : : ||| ||||| ||||| ||||| ||||| : ||||| : ||||| :
Db 17 ISCNNSG--GDSASTNPDESAKGNPLVTSKRTTDSNAVLAVKEVALSSIDEI--SKA 73
   :||||| : : : ||| ||||| ||||| ||||| ||||| : ||||| : ||||| :
QY 61 IGGKIHONNGLDTEYHNHNSLAGAYAISTLIKQKIDGLKNELGKKEKIDAAKCSSEFTN 120
   |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 74 IGGKIKNDGTLDNANNNESLAGAYEISKLITQKLSVNSELEKKEIKKAKKCSSEFTT 133
   |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 121 KLEKHTDLGKEGVTDADAKFAILKTNGTKGAEELGKFESEVYLSKAKKEMLAN 180
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 134 KLDNSHAELGIGQVDDNKKALIKTHGKDKAKLEBELFESLSLSAAQAALTNVSK 193
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 181 ELTSPVVAESP 191
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 194 ELTNPVVAESP 204
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

```

RESULT 4
US-09-708-427-19883
Sequence 19883, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

```

```

Query Match          9.5%; Score 175; DB 6; Length 1014;
Best Local Similarity 24.3%; Pred. No. 0.00099;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

```

```

QY 8 KDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKE 45
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 401 KHGETEADSKGYLGVAELQSTLEAFQVKSSSLEALINATENKELTENLVNTESEKK 460
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 46 VEALLSSIDEIAK-----AIGKIHONNGLDTEYHNHNSLAGAYAISTL--IKQ 94
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 461 LEA---TVDEYSVAKISESENLISIRELNVTOG--KLESTIENDLKAAGLOESEVMEKLS 516
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 95 KLDGLKNEGLKEKIDAA-----KKCSSEFTNK-----LKEKH 126
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 517 AEEISLEQKG--REIDEATTKRMELALHQSLSIDSEHRLOKAMEEFTSRDSEASSLTELK 574
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 127 TDL--GKEGVTDADAKFAILKTNGTKGAEELGKFESEVYLSKAKKEMLANVKELTSP 185
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 575 RDLGKIKISYEQLAASGRSSSLKLEKLEQTLGRLAASVNEKLEQEPDQAQESLGSS 634
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 186 VVAESPAMVNNSGK-----DGNTSANSADSVKGNLTETSKRTTDSNAVLAVKE 235
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 635 SESELLAETNNOLAKIKIOLEGLIGSGVKEKALKRLAEALERNNOKEETESSDIVERIK 694
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 236 EVELTSSIDEIA--KAIG-----KKIKN----- 257
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 695 THENOIEYRKKLAHSAQVADTRKVELDALSKIKLESTIEBELGAKCOGLEKESGDLAE 754
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 258 -DVSLDNEADHNGSLISGAYLISNLTTRKISAIKNSGELKA--ETEKAKKCSSEFTAKL-- 313
   :|| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

```
Db 755 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELASAKTTIEDTLKQLTS 807
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVK-NLSKAKEMLTNSVK- 371
Db 808 EGEKLOSOIEKLRVAVAEKSVSLESH-----PEELEKTLSEVKAQLKEVNAATAVSVKV 861
QY 372 -ELTS 375
Db 862 AELTS 866

RESULT 5
US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match 9.5%; Score 175; DB 6; Length 1018;
Best Local Similarity 24.3%; Pred. No. 0.0004;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

QY 8 KDGNTSANSAD-----SVKGPNL-----TEISKRTDS-NAVLAVKE 45
Db 405 KHGETEADSKGYLGOVAELQSTLEAFQVKKSSLEALNATENEKELETLNNAVTSSEKK 464
QY 46 VEALLSIDELIAK-----AIGKKIHONGLDTEYNNHNSLAGAVASTL--IKO 94
Db 465 LEA--TVDEYSVKISESENLESINRELNTVG--KLESIENDLKAAGIOESEVMEKLS 520
QY 95 KLDGLKNEGLKEKIDNA-----KCSFTPTNK-----LKEKH 126
Db 521 AEESELEQK--REIDEATTKRMELALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 578
QY 127 TDL-GKEGYTDADAKAAILKTNGTKGAPELGLFESEVEVLSKAKEMLANSVRELTP 185
Db 579 RDLERKITSYEQDLAEASGSSSLKELEQTLGRLLAAESVNEKLEKQEDQAOEKSLOSS 638
QY 186 VVAESPAMVNNNGK-----DGNTSANSADSEVKGPNLTE--ISKKTESNAVAVLAK 235
Db 639 SESELLAETNNQKTKIKIQLEGLIGSGVEKETALKRLEAIFERNQKTESSDVLEKIK 698
QY 236 EVELTLTSDILA-KAIG-----KKIKN----- 257
Db 699 THENOIEEYKRLAHEASGVADTRKVELDALSKLKNLESTIEELGACOGLEKESGDLAE 758
QY 258 -DVSLDNEADHNGSLISGAYLISNLTTRKISAKDSGEKLA-EIEKAKCSEFTAKL-- 313
Db 759 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELASAKTTIEDTLKQLTS 811
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVK-NLSKAKEMLTNSVK- 371
Db 812 EGEKLOSOIEKLRVAVAEKSVSLESH-----PEELEKTLSEVKAQLKEVNAATAVSVKV 865
```

```
QY 372 -ELTS 375
Db 866 AELTS 870

RESULT 6
US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881
```

```
Query Match 9.5%; Score 175; DB 6; Length 1269;
Best Local Similarity 24.3%; Pred. No. 0.00053;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;
```

```
QY 8 KDGNTSANSAD-----SVKGPNL-----TEISKRTDS-NAVLAVKE 45
Db 656 KHGETEADSKGYLGOVAELQSTLEAFQVKKSSLEALNATENEKELETLNNAVTSSEKK 715
QY 46 VEALLSIDELIAK-----AIGKKIHONGLDTEYNNHNSLAGAVASTL--IKO 94
Db 716 LEA--TVDEYSVKISESENLESINRELNTVG--KLESIENDLKAAGIOESEVMEKLS 771
QY 95 KLDGLKNEGLKEKIDNA-----KCSFTPTNK-----LKEKH 126
Db 772 AEESELEQK--REIDEATTKRMELALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 829
QY 127 TDL-GKEGYTDADAKAAILKTNGTKGAPELGLFESEVEVLSKAKEMLANSVRELTP 185
Db 830 RDLERKITSYEQDLAEASGSSSLKELEQTLGRLLAAESVNEKLEKQEDQAOEKSLOSS 889
QY 186 VVAESPAMVNNNGK-----DGNTSANSADSEVKGPNLTE--ISKKTESNAVAVLAK 235
Db 890 SESELLAETNNQKTKIKIQLEGLIGSGVEKETALKRLEAIFERNQKTESSDVLEKIK 949
QY 236 EVELTLTSDILA-KAIG-----KKIKN----- 257
Db 950 THENOIEEYKRLAHEASGVADTRKVELDALSKLKNLESTIEELGACOGLEKESGDLAE 1009
QY 258 -DVSLDNEADHNGSLISGAYLISNLTTRKISAKDSGEKLA-EIEKAKCSEFTAKL-- 313
Db 1010 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELASAKTTIEDTLKQLTS 1062
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVK-NLSKAKEMLTNSVK- 371
Db 1063 EGEKLOSOIEKLRVAVAEKSVSLESH-----PEELEKTLSEVKAQLKEVNAATAVSVKV 1116
QY 372 -ELTS 375
Db 1117 AELTS 1121
```

RESULT 7

```

US-09-708-427-15044
; Sequence 15044, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15044
; LENGTH: 1313
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1313
; OTHER INFORMATION: Xaa is any amino acid
; OTHER INFORMATION: Ceres Seq. ID 1828627
US-09-708-427-15044

```

```

Query Match
Best Local Similarity 9.2%; Score 169.5; DB 6; Length 1313;
Matches 118; Conservative 60; Mismatches 140; Indels 213; Gaps 23;

QY 8 KDGNTSANSADSVKGNPNTLEISKRTTDSNAV-----LAVKEVALLSIDEIATAKA 60
DB 6 KTGIMETPRKSPPPRRLSKLSASKSDNSASPKVHSRLVKGTGLQTLNOIOEDLKA 65
QY 61 IGGKHQNGNGLDTEYVNHGSLAGAVASTLTKOK-----LDGLK-----NEGIEKI 108
DB 66 -----DEQ-----LELLKKAKAIDDLKESELYEANEKIKLEAL 101
QY 109 DAARKCSETF-----TNK-----LKEKHT-DLG----- 130
DB 102 AAQKAEESEFEYKFRVLEQAGLEAVOKKDYTSKNELESIRSGHALDISALSTTEEL 161
QY 131 -----KEGYTDADAEKAILKTGTYTGAE-----ELGKL-----FES 163
DB 162 QVKEHELSMTADAKNKAISHAEATKTAETHAKAEILASLEGRLKALLGSKKEEAL 221
QY 164 VEVSRAKE-----MLANSVKE-----LTSFVVAESPAMVNSG-- 198
DB 222 MEIVSKLSELELKGELKEVYSITESLKEQGLVBOLKVDLEAKAKAES-----CTNSV 278
QY 199 -----KDGNTSANSADSVKGNPNTLEISKRTTESNAV----- 231
DB 279 EMNKVHELEKEVEESNRKSSASMS-----SYMKOLAEHLNHYLHETKSDNAQKEKI 333
QY 232 -LAVKEVETLTSIDE-----LARAIGKIKNDY-----SIDNEADHN 268
DB 334 ELEKTEIEMQRTDLEERYGQVCIACEEASKLENLESIKSELEISOEKTTRALDNEKAT 393
QY 269 GSILISAVYIISMLITTKISAIKDSGELKAELEKAKKSEFTAKLKGERTDLKREGVTD 328
DB 394 SN-----IQMLDQFELSELERCKVEBEKSKDMESTLAL-----QENASTES 438
QY 329 NAKKAILTNNDKRGADLEKLEFESYKNSKAKE-----MLTNSYKELTS 375
DB 439 SEAKATLLVCOBEIKNC-----ESQVDSIKLASKFTNEKYEMKLEADANEIDS 486

```

```

RESULT 8
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone

```

```

; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRF
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

```

```

Query Match
Best Local Similarity 9.2%; Score 169; DB 7; Length 1881;
Matches 98; Conservative 79; Mismatches 172; Indels 124; Gaps 16;

QY 1 MACNNSGKDGNTSANSADSVKGNPNTLEISKRTTDSNAVILAVKEVALLSIDEIATAKA 60
DB 1411 LAAQLETSTSHDSTTWELEKTELEKRVNELTEATSELTQLDNNOSLIELEKTKAA 1470
QY 61 IGGKHQNGNGLDTEYVNHGSLAGAVASTLTKOKIDGLKNE-----GLKEKIDAAK 114
DB 1471 LTK-----SSKLEVGCGNOKSELSDPSLKS-----VKSELKNEKNYQNETTSLKDETEER 1523
QY 115 SEFTFNKLEKHEHDLKEGVTADAKAEALKT-----NG 148
DB 1524 IVTLQTELEKDRISEVEKEEAMSENSETVIKEYSDKIKLSKINSIKENHSEITTHNE 1583
QY 149 TKTGAEELEKLEFESYVLSKAKEMLANSYKELTSPVVAESPAMVNSGKDGNTSANS 208
DB 1584 QKTSLKQDIATKLSQHE--SAQTOLEDEKNOELKASL-----EKHNTESATSI 1631
QY 209 DESVKGPNLTESKRT-----TESNAVYLAKEVETLTSIDELARAIGK-----K 254
DB 1632 EE--RNNOIKELSETIKSLKTELTSQDALKOSKEKTKTKNSDPTESKLEKQEELEK 1669
QY 255 IKNDVSLDNEADHNGSLISGAYIISMLITTKISAIKDSGELKAE-----ELEKAK-- 303
DB 1690 VKSDIQRTADE-KLKGITERETALKSELEYKNSGLSTTSELAALTKYKSLKEKEKELOF 1748
QY 304 -----KCSSEFTAKLKGERTDLGK--EGVTD-----DNAKAILKTNNDKRGAD 346
DB 1749 LSGKSKLEDEDYLOK-----HSDISEKTKALIDELKEKTKQFDSDSKKLTLELNDLSTTK 1804
QY 347 ELE-----KLFEVKNLSKAKEMLTNSYKELTS 375
DB 1805 ELEETKQTSKFKNLEBKDKETIVLKNLELKNLNDNSAKKEISEKVS 1857

```

```

RESULT 9
; Sequence 15590, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Itawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard

```

```

0Y      1 MACNNGKCGKGNFSANSAADSVKGNPLTEISKRTDGSNAVLAVKEAYEALSSIDEIATAKA 60
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1411 LAAOLETKTSNLDSTTMELEKTELKVKVRELTETSELTIKLQDNNSLTLEIEKTAKA 1470
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      61 ICKKIHNNGGLDTEVYNHNGSLLAGAYAISTLIKOKLDGLKNE-----GIKEKITDAKKC 114
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1471 LKR-----SKMDLEVCGNQKSELQDSLS-----VKSLEKNENKXNQETTSLKDEIEEKQE 1523
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      115 SETFTNKLEKKEHTDLCEGVTDADAKEALIKT-----NG 148
        :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1524 IYTLOTELKDRISVEKERAMLSENSEFTVIKEYSDKIKSLESKINSIKEHNSKEITTHNE 1583
        :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      149 TTTKGAEELGRPESEVFNLSKAKEMLASVKELTSPYAESAAMNNSKDGTNSANSA 208
        :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1584 OKTSCLKODIAKSODHE-SAQTOLEDKENOLAKELKASL-----BKHNTESTATSI 1631
        :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      209 DRSVKGPNLTEISKRI-----TESNAVFLAVKEVETLLTIDELAKAIKGR-----K 254
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1632 EE--KNNOJKELSEITSLSKLTELKTSGDALKQSOKREYKTLKTNSDPESLEKOLELEK 1689
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      255 INDVSLNDREADHNGSLISGAYLISNLITRKISAIKDSGELKA-----BIEKAK----- 303
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1690 VKSDLOTADE-KLKGITEREIALKSELETIVKNSGLSTTSBLALTQTVKSELEKEBELQF 1748
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      304 -----KCSEETAKUKGGHTDLGK--EGVTD-----DNAKKAILKTNDDKTGAD 346
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        1749 LSGNKSKELEDYIOK---HSDISERKLKALTDLEKKEKTFOPDSSKKRLTLELNDITSTKK 1804
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
0Y      347 ELE-----KLFEYSKNISSKAKEMLTNSVKELTSS 375
        :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1805 ELETEKTOTSKRKNLEERRDKETIVKLNKELELKNDSNGAKKELSEKVSKELES 1857
        :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 10
US-09-708-427-15045
; Sequence 15045, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15045
; LENGTH: 1304
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

```

Oy	24	PNTEISIKRIDSNANV-----	FLAVEYALLSIDETIAKAIKKIKHONNGLDEYN	76	
		: : :   :   :	: : :   :   :		
Dd	13	PRLSLTSKSDSDNSASPKVSRLYKGELOTOLONOEDLKA	-----	-DQ-- 59	
Oy	77	HNGSLLAGAVAISTLIKOR--	LDGLK-----	-NEGLEKIDAAKCSETF----- 118	
		: : :   :   :	: : :   :   :		
Dd	60	-----	IELLKOKAKAIDDLKESEKLEVEANBKLEALAAQKRAESESVEFKR	108	
Oy	119	-----	-TNK-----	LKEKHT-DLG-----	-KGCYTDADAKEA 142
		: :   : :   :	: :   : :   :	: :   : :	
Dd	109	AVELEOAGLEAVOKKDYTSKNELESIRSQHALDISALLSTTEELQRYKHLSHMTADAKNK	168		
Oy	143	ILKTNGTTRKAE-----	ELCKL-----	FEEVELSKAAE-----	173
		: :   :   :	: :   :   :	: :   :   :	
Dd	169	ALSHAPEATKLAETIHAERKAETILASELGLKALLASKKEKEAIEGNETIVSKLSKSEIELLRG	228		
Oy	174	-----	-MLANSVKE-----	LTPVAAESPPAMVNNSG-----	KDG 201
		: :   :   :	: :   :   :	: :   :   :	
Dd	229	ELEKYSILLESKIEKEDGLEVOLKYDLEAKKAABES--	CTNSVSEEMKKNKHYELEKEVEBS	265	
Oy	202	NTSANSADESVKGPMLTEISIKRITESNANV-----	-----	-LAVKEVELTTSIDE 246	
		: :   :   :	: :   :   :	: :   :   :	
Dd	286	NRSKSSASESME-----	-SYMQLAMELNVHETKSDNAQOEKTELEKTEIQAQRTDLEE	340	
Oy	247	-----	-LAKAIGKTRKNDV-----	-----	SLDNEADHNGSLISCAVILSNLTK 284
		:   : :   :   :	:   : :   :   :	:   : :   :   :	
Dd	341	YGRQVCIAKEKASLENIENVESIKSELEISOEKTRALDNEAATSN-----	-----	-TONLLDQ 393	
Oy	285	KISATKDSGELKAETIERAKKCSSEPTAKLKEBHNDLKGEGYTDONAKKAILKTNDDTKG	344		
		: : :   :   :   :   :	: : :   :   :   :   :	: : :   :   :   :   :	
Dd	394	RTELSIELERCKVEEKSCKDMESITLAL-----	-----	-OEAATESEAKATLLVCQEBELKN	445
Oy	345	ADELEKLFESYKNLSKAKE-----	MLTNSVKEIYS 375		
		: :   :   :	: :   :   :	: :   :   :	
Dd	446	C---	ESQYDSTLKLASKETNEKTEKMLEBARNEIDS 477		

```

RESULT 11
US-09-748-875-14
; Sequence 14, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 589
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-14

```

Query Match 8.6%; Score 158; DB 5; Length 589;

Best Local Similarity 25.1%; Pred. No. 0.0028;  
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 1 EGVSGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 51
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLGLKNEGJKEKIDA 110
Db 52 LITKISEIKKKYLDLKVNVLSAEELTSKTEKTEKELTATPEOKKDTLPTEPEKVAEA 111
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 112 OKKVEAKKKAEDQEKDRNRPITYKTLLELAESDVEVKAELDELIVAKKESQDEE 171
OY 159 KLFES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190
Db 172 KIKOAEAVESKQAEATRLKIKTDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 231
191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNAVVLAVKEVELTTS 243
Db 232 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KDVDEYI-- 281
OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISNLTKKISAIKDSGELKAEIEKAK 303
Db 282 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 322
OY 304 KCSEFTAKLKGHTDLGEGVTDNNAKAILKTNDKTKGADELKLPESVKNLSKAAK 363
Db 323 SKKEELTSKTKAELT-----AAFEQFKDTLKPE-----KKVAEAKKVEAKKRAKADOK 372
OY 364 E-----MLTNSVKEL 373
Db 373 EEDRRNYPNTYKTL 387
```

RESULT 12  
US-09-748-875-61

; Sequence 61, Application US/09748875  
; GENERAL INFORMATION:

; APPLICANT: BRILES et al.

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-61

Query Match 8.6%; Score 158; DB 5; Length 690;  
Best Local Similarity 25.1%; Pred. No. 0.0035;  
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 38 EGVSGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 88
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLGLKNEGJKEKIDA 110
Db 89 LITKISEIKKKYLDLKVNVLSAEELTSKTEKTEKELTATPEOKKDTLPTEPEKVAEA 148
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 149 OKKVEAKKKAEDQEKDRNRPITYKTLLELAESDVEVKAELDELIVAKKESQDEE 208
```

```
OY 159 KLFES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190
Db 209 KIKOAEAVESKQAEATRLKIKTDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 268
OY 191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNAVVLAVKEVELTTS 243
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KDVDEYI-- 318
OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISNLTKKISAIKDSGELKAEIEKAK 303
Db 319 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 359
OY 304 KCSEFTAKLKGHTDLGEGVTDNNAKAILKTNDKTKGADELKLPESVKNLSKAAK 363
Db 360 SKKEELTSKTKAELT-----AAFEQFKDTLKPE-----KKVAEAKKVEAKKRAKADOK 409
OY 364 E-----MLTNSVKEL 373
Db 410 EEDRRNYPNTYKTL 424
```

RESULT 13  
US-09-748-875-1

; Sequence 1, Application US/09748875  
; GENERAL INFORMATION:

; APPLICANT: BRILES et al.

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-1

Query Match 8.6%; Score 158; DB 5; Length 691;  
Best Local Similarity 25.1%; Pred. No. 0.0035;  
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 38 EGVSGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 88
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLGLKNEGJKEKIDA 110
Db 89 LITKISEIKKKYLDLKVNVLSAEELTSKTEKTEKELTATPEOKKDTLPTEPEKVAEA 148
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 149 OKKVEAKKKAEDQEKDRNRPITYKTLLELAESDVEVKAELDELIVAKKESQDEE 208
OY 159 KLFES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190
Db 209 KIKOAEAVESKQAEATRLKIKTDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 268
OY 191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNAVVLAVKEVELTTS 243
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KDVDEYI-- 318
OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISNLTKKISAIKDSGELKAEIEKAK 303
Db 319 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 359
OY 304 KCSEFTAKLKGHTDLGEGVTDNNAKAILKTNDKTKGADELKLPESVKNLSKAAK 363
Db 360 SKKEELTSKTKAELT-----AAFEQFKDTLKPE-----KKVAEAKKVEAKKRAKADOK 409
```



```
QY      364 E-----MLTNSVKEL 373
          |      ||: | |
Db      410 EEDRRNYPNTYKTL 424
```

RESULT 14  
US-09-748-875-62

```

: Sequence 62 Application US/09748875
:
: GENERAL INFORMATION:
:
: APPLICANT: BRILES et al.
:
: TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
:
: TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
:
: FILE REFERENCE: 454312-3140
:
: CURRENT APPLICATION NUMBER: US/09/748,875
:
: PRIOR FILING DATE: 2000-12-26
:
: PRIOR APPLICATION NUMBER: 09/298,523
:
: PRIOR FILING DATE: 1999-04-23
:
: NUMBER OF SEQ ID NOS: 78
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 62
:
: LENGTH: 701
:
: TYPE: PRT
:
: ORGANISM: Streptococcus pneumoniae
:
: US-09-748-875-62

```

Query Match	8.6%	Score 158:	DB 5:	Length 701:
Best Local Similarity	25.1%	Pred. NO. 0.0035:		
Matches 109; Conservative	49;	Mismatches 149;	Indels 128;	Gaps 20

```

0Y 19 ESUV-GPULT-----EISKHTDSNNAVLAVKEVEALLSIDIELAKKAIGKIKH-ONNG 70
Db 38 EGVSNANLIVTSSGQDSSKRYAD-----EVESHLESILKDVKKNLKKKVOHONTNG 88
0Y 71 LDTE-----YNHNGSLAGAYAIS-----TLIKOKLDGLNEGLEKIDA 110
Db 89 LITRKEISKKKYLVLDKNNVLSBAELNSKTEKTELTAPEQKKOTLPTEPEKKVAEA 148
0Y 111 AKKSEFTTNKLEKEND-----LGE-GVTDADAKAELKTNGTKTKGAELG 158
Db 149 OKKVEAKKKAEDQEKEDRRNYPITYTKTLELTAESDVEYKKAELVLVYVKAKEODEE 208
0Y 159 KLFPES-VEVLAK-----AAKEMLANSVKLELSPVVAES----- 190
Db 209 KIKQAEAEVESKQAEATRLKTIKTIDREAKRKADAKLAEAVEKKVATSEDDCKPRRAKRG 268
0Y 191 -----PAMVNSGCKDGNTSANSADESVKGNLTETSKITTESNAVLAVKEVELTTS 243
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR-----KVDDEXI-- 318
0Y 244 IDELAKAKGKIKINDVSLDNADNINGSLSISAYLSNLITRKISAIADSGELAKAEIFKAK 303
Db 319 -----KKMLSEIQTDR-KHTQAV-----NL-NIKISAIKTYLYELTELSTYKEN 359
0Y 304 KCSSEFTAKLGEHTDLCGEVTDONDKAKKALIKTNNDKTGADAELEKLFESVKNLSKAAK 363
Db 360 SKKEELTSKTKAEALT-----AAPEQFKOTLKE-----KKVAEAKKVEAKKKAADOK 409
0Y 364 E-----MLTNSVKEL 373
Db 410 EEDRRNYPNTNYKTL 424

```

```

RESULT 15
US-09-748-875-2
; Sequence 2, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et. al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 45312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875

```

: CURRENT FILING DATE: 2000-12-26  
 : PRIOR APPLICATION NUMBER: 09/298,523  
 : PRIOR FILING DATE: 1999-04-23  
 : NUMBER OF SEQ ID NOS: 78  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 2  
 : :  
 : LENGTH: 707  
 : TYPE: prt  
 : :  
 : ORGANISM: Streptococcus pneumoniae  
 :  
 : US-09-748-875-2

Query Match	8.6%	Score 158;	DB 5;	Length 707;
Best Local Similarity	25.1%	Pred. No. 0.0036;		
Matches 109; Conservative	49;	Mismatches 149;	Indels 128;	Gaps 20;

Oy	19	ESV- GPHLT-----E-SKRTITSONAVLAVKEVALLESIDELIAKAIGKRH--ONG 70
Dd	38	EGBRGNNLTITSSGQDISKRYAD-----EVESHLESILKDVKYNNLKRYQHONMG 88
Oy	71	LDT-----YNHGSLLAGAYAIS-----TLIKOKLDGLNEGLEKIDA 110
Dd	89	LITRSEIKRKYYLDLKNVLSEAFELTSKTKETKEKLTAPEOPKKDPTLPTEPERKAAEA 148
Oy	111	AKKCEFTTNMLKEHNH-----LGKE-CGYDDADAKELILKTNGTGTGAEBLG 158
Dd	149	OKKVEAARKKADOKEKDRNYPITYYKTLELELAEDSVEVKKAEVLVYKVAKESODEE 208
Oy	159	KLFES--VEVLK-----AAKEMLANSVKELTSPPVAAES-----190
Dd	209	KIKOALEVESKOAEATPLKIKITTDREARAKKADAKILEANENKVANISBODKPRRAKRG 266
Oy	191	-----PAMVNSGKGDNITSANSADESVKCPNLTEISKTTSESNAVYLAVEYTELLTS 243
Dd	269	VSGELATPDCKKENAPKSSDSSEV--GEELTPSPSMANESOTERR-----KDVDEYI- 318
Oy	244	IDELAKAI GKIRKNDVSJDNEDAHNGSLISAGAYLSNLTIRKISAIRKDSGELAIEIKAK 303
Dd	319	-----KKMSELIOUDGR-KITPNV-----NL-NIKLSAITTKLYILELSTVLKEN 355
Oy	304	KCSEEF TAKLGEHTDLQEGVTDDNAKKAILKTNNDKTGACADELEKLFESVKNLSAAK 363
Dd	360	SKKEELT SKTKAEFLT-----AAFEOPFKDPTLKPE-----KKVAABEKKVEBAKKKAKDOK 409
Oy	364	E-----MTNSVKEL 373
Dd	410	FEDRRNYPTNTYKTL 424

Search completed: March 18, 2002, 09:58:16  
Job time: 374 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:37 ; Search time 68.77 Seconds  
(without alignments)  
415.377 Million cell updates/sec

Title: US-09-596-746A-24

Perfect score: 1842  
Sequence: 1 MACNNSGKDCDNTSANSNADES.....KNLSKAKKEMLTNSVKELTS 375

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR.68:\*

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	50.4	210	2	G70218
2	921	50.0	210	2	S69927
3	879	47.7	210	2	I40144
4	853	46.3	177	2	I40129
5	757.5	41.1	211	2	I40145
6	756.5	41.1	193	2	S70279
7	696.5	37.8	193	2	S70287
8	683.5	37.1	211	2	I40277
9	677.5	36.8	211	2	S69918
10	676.5	36.7	209	2	I40273
11	674.5	36.6	191	2	S70278
12	674.5	36.6	191	2	I40153
13	673.5	36.6	211	2	I40268
14	672.5	36.5	193	2	S70280
15	666	36.2	194	2	S70277
16	665.5	36.1	209	2	S69917
17	665	36.1	190	2	S70273
18	661	35.9	212	2	S70254
19	657.5	35.7	193	2	S70276
20	655.5	35.6	193	2	S70274
21	654.5	35.5	211	2	I40278
22	654	35.5	212	2	S69922
23	653.5	35.5	191	2	S70284
24	652.5	35.4	193	2	S70265
25	652.5	35.4	211	2	S69932
26	652	35.4	194	2	S70268
27	646	35.1	212	2	I40279
28	642.5	34.9	209	2	I40142
29	635	34.5	212	2	I40143

30	634.5	34.4	203	2	I40108	outer surface prot
31	632	34.3	214	2	S69916	outer surface prot
32	631.5	34.3	193	2	S70286	outer surface prot
33	629.5	34.2	191	2	S70288	outer surface prot
34	626	34.0	194	2	S70289	outer surface prot
35	624.5	33.9	179	2	S54222	outer surface prot
36	623.5	33.8	209	2	I40285	outer surface prot
37	621.5	33.7	209	2	I40281	outer surface prot
38	620.5	33.7	178	2	I40125	outer surface prot
39	619	33.6	210	2	S69925	outer surface prot
40	618	33.6	192	2	S70285	outer surface prot
41	617.5	33.5	178	2	I40104	outer surface prot
42	615	33.4	210	2	S69923	outer surface prot
43	614	33.3	212	2	S20543	outer surface prot
44	613	33.3	210	2	S69920	outer surface prot
45	610	33.1	210	2	I40272	outer surface prot

## ALIGNMENTS

RESULT 1  
G70218  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence:Revision 13-Feb-1998 #extl:change 21-Jul-2000  
C:Accession: G70218; I40269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kevlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: G70218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <KLE>  
A:Cross-References: GB:AE000792; NID:g3253098; PIDN:AA06329.1; PID:g2689901; TIGR:BB  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu  
A:Reference number: I40269; MUID:96025162  
A:Accession: I40269  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-References: GB:D49497; NID:g707092; PIDN:BA08457.1; PID:g769684  
R:Jauris-Helpe, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os  
A:Reference number: S37726; MUID:93268136  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <TAU>  
A:Cross-References: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dornier, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
A:Reference number: S70255; MUID:9626448  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-References: EMBL:I42887; NID:g858715; PIDN:AA03695.1; PID:g1695212  
A:Experimental source: strain Ip2  
C:Genetics:  
A:Gene: ospC  
A:Superfamily: Lyme disease spirochete surface protein C

## Query Match

Best Local Similarity 50.4%; Score 928; DB 2; Length 210;  
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60  
QY 61 ICKKTHONNGDTEYNHNSGLAGAVASTLIKOKLDGKNGLEKIDAAKCKSEFTFN 76  
DB 77 ICKKTHONNGDTEYNHNSGLAGAVASTLIKOKLDGKNGLEKIDAAKCKSEFTFN 120  
QY 121 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180  
DB 137 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180  
QY 181 ELTSPVAESP 191  
DB 197 ELTSPVAESP 207

## RESULT 2

S69927

outer surface protein C precursor - Lyme disease spirochete (strain PKa)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69927; S72669

R:Jauris-Heipke, S.; Liegl, G.; Preac-Murisc, V.; Roessler, D.; Schwab, E.; Soutschek, H.

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia

A:Reference number: 140047; MID:95395018

A:Accession: S69927

A:Status: nucleic acid sequence not shown

A:Residues: 1-210 <JAU>

A:Cross-references: EMBL:X69589

R:Jauris, S.

submitted to the EMBL Data Library, February 1994

A:Accession: S72669

A:Molecule type: DNA

A:Residues: 1-124, 'D', '126-139', 'E', '141-210 <JAM>

A:Cross-references: EMBL:X69589

C:Genetics: none: ospC

perfamily: Lyme disease spirochete surface protein C

## Query Match

Best Local Similarity 50.0%; Score 921; DB 2; Length 210;  
Matches 186; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60  
QY 61 ICKKTHONNGDTEYNHNSGLAGAVASTLIKOKLDGKNGLEKIDAAKCKSEFTFN 76  
DB 77 ICKKTHONNGDTEYNHNSGLAGAVASTLIKOKLDGKNGLEKIDAAKCKSEFTFN 120  
QY 121 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180  
DB 137 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180  
QY 181 ELTSPVAESP 191  
DB 197 ELTSPVAESP 207

## RESULT 3

I40144

outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40129; S54199  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

OY 11 NTSANSADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAAKIGKIHONNG 70  
 Db 1 NTSANSADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAAKIGKIHONNG 60  
 OY 71 LDTENVHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNKLEKHTDLC 130  
 Db 61 LDTENVHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNKLEKHTDLC 120  
 OY 131 KEGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNVKELTSPVV 187  
 Db 121 KRQVTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNVKELTSPVV 177

RESULT 5  
 140145  
 outer surface protein C precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 C:Accession: I40145  
 R:Stevenson, B.; Barthold, S.W.  
 Mol. Microbiol. Lett. 124, 367-372, 1994  
 A:Title: Expression and sequence of outer surface protein C among North American isolates  
 A:Reference number: I40143; MUID:95154673  
 A:Accession: I40145  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-211 <RES>  
 A:Cross-References: EMBL:U04282; NID:g2314881; PIDN:AA045540.1; PID:g434666  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.1%; Score 757.5; DB 2; Length 211;  
 Best Local Similarity 81.2%; Pred. No. 1.1e-28;  
 Matches 156; Conservative 13; Mismatches 22; Indels 1; Gaps 1;  
 OY 1 MACNNSGKCGN-TSANSADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAK 59  
 Db 17 ISCNNSGKCGNMASTNPADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAK 76  
 OY 60 AICGKIHNNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFT 119  
 Db 77 AICGKIHNNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFT 136  
 OY 120 NKLKHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 179  
 Db 137 NKLKHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 196  
 OY 180 KELTSPVVAESP 191  
 Db 197 KELTSPVVAESP 208

RESULT 6  
 S70279  
 outer surface protein C - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70279  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70279  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIV>  
 A:Cross-References: EMBL:L42898; NID:g858729; PIDN:AA037007.1; PID:g1695223  
 A:Experimental source: strain 25015  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.1%; Score 756.5; DB 2; Length 193;  
 Best Local Similarity 82.1%; Pred. No. 1.1e-28;  
 Matches 156; Conservative 10; Mismatches 23; Indels 1; Gaps 1;  
 OY 3 CANSNGKCGN-TSANSADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAKAI 61  
 Db 1 CANSNGKCGNMASTNPADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAKAI 60  
 OY 62 GKRIHONNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNK 121  
 Db 61 GKRIHONNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNK 120  
 OY 122 LKEKHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 181  
 Db 121 LKSHHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 180  
 OY 182 LTSPVVAESP 191  
 Db 181 LTSPVVAESP 190

RESULT 7  
 S70287  
 outer surface protein C - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70287  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70287  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIV>  
 A:Cross-References: EMBL:L42895; NID:g858723; PIDN:AA037003.1; PID:g1695220  
 A:Experimental source: strain 28354  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.8%; Score 696.5; DB 2; Length 193;  
 Best Local Similarity 76.3%; Pred. No. 6.4e-26;  
 Matches 145; Conservative 16; Mismatches 28; Indels 1; Gaps 1;  
 OY 3 CANSNGKCGN-TSANSADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAKAI 62  
 Db 1 CANSNGKCGNMASTNPADSVKGPNTLTSKRTTDSNAVLAVKEVEALLSTIDEIAKAI 60  
 OY 63 KRIHONNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNK 121  
 Db 61 KRIHONNGLDTEYNHNGSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSSEFTFNK 120  
 OY 122 LKEKHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 181  
 Db 121 LKSHHTDLGKGYTDADAKKALKTNGTKTGABELGKLFESVEVLSKAEMLSNV 180  
 OY 182 LTSPVVAESP 191  
 Db 181 LTSPVVAESP 190

RESULT 8  
 I40277  
 outer surface protein C precursor - Borrelia afzelii  
 C:Species: Borrelia afzelii  
 C>Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
 C:Accession: I40277  
 R:Fukunaga, M.; Hamase, A.  
 J. Clin. Microbiol. 33, 2415-2420, 1995

A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato  
A:Reference number: 140259; MUID:96025162  
A:Accession: 140277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-211 <RES>  
A:Cross-references: GB:D49501; NID:g707096; PIDN:BA09461.1; PID:g769688  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match	37.1%	Score 683.5;	DB 2;	Length 211;
Best Local Similarity	75.1%	Pred. No. 2.8e-25;		
Matches 145;	Conservative 19;	Mismatches 35;		

[illegible]

RESULT 9  
 569918  
 Outer surface protein C precursor - Lyme disease spirochete (strain PBr)  
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
 A:Variety: Strain PBr  
 C:Date: 06-Dec-1996 #sequence-revision 14-Feb-1997 #text-change 26-May-2000  
 C:Accession: S69918; S72674; I40103  
 R:Jauris-Heldke, S.; Liegl, G.; Preacg-Hurduc, V.; Roessler, D.; Schwab, E.; Soutschek, F.  
 J. Clin. Microbiol. 33, 1860-1866, 1995  
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia burgdorferi*  
 A:Reference number: 140047; MUID:95395018  
 A:Accession: S69918  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-211 <JAU>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022  
 Experimental source: Strain PBr  
 Roessler, D.  
 Submitted to the EMBL Data Library, September 1994  
 A:Reference number: S72674  
 A:Accession: S72674  
 A:Molecule type: DNA  
 A:Residues: 1-152, 'E', 154-211 <ROE>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

	Query Match	36.88;	Score 677.5;	DB 2;	Length 211;
Best Local Similarity	77.28;		Pred. No. 5.3e-25;		
Matches 149;	Conservative 11;	Mismatches 30;	Indels 3;	Gaps 3;	
QY	1	MACNNSGKDGNTSANSADSVKGPNTLEISKRTITPSNVLAVVEVALLSSIDEIAAKA	60		
		::			
Db	17	ISCNNSGKDGNTSANSADSVKGPNTLEISKRTITPSNVLAVVEVALLSSIDEIAAKA	75		
		::			
QY	61	IGKRIHONNGLDREYHNHNSLLAGAAVSTLIKQKLDGLK-NEGIKRIIDAAKCESETFT	119		
		:			
Db	76	IGKRIKRDGSLGPEAHNHNSLLAGAAVTTSTLTQKLSLNSEGLKRIIAAKKCEEFS	135		
		:			
QY	120	NKLEKHHTDIGKRGVYTDADAKEAIIKTNGT-KTNGAEELIGLFESEVYLSKAAKEHLANS	178		
		:			

Db 136 TKLDNNAOGCTGGTATDDNNAKKAIIKANAAGCKDGYEELEKJLSLESTSKAÅEMLANS 199

QY 179 VKELTSPVVAESP 191

Db 196 VKELTSPVVAESP 208

```

RESULT 10
I40273
outer surface protein C precursor - Borrelia afzelii
C:Species: Borrelia afzelii
C:date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C:Accession: I40273
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; MUID:96025162
A:Accession: I40273
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-209 <RES>
A:Cross-references: GB:D49379; NID:g1041107; P1DN:BA08377.1; P1D:g1041108
C:Superfamily: Lyme disease spirochete surface protein C

```

	Query Match	36.7%	Score 676.5	D8.2f	Length 209;
	Best Local Similarity	75.0%	Pred No. 5,9e 25;		
	Matches 144;	Conservative 20;	Mismatches 25;	Indels 3;	Gaps 3;
Oy	1 MACNNSGCDGTSTNSADSESVKQPNPTLTSKITDTSNAVLAAKEVALLSIDEIAANA	60	:::     ::: :::		
D8	17 ISCNNSGGKDGIASITNPDESAGKPULTEISKRTIDTSNAVLAIVEVALLSIDEI-AAAT	75	:::     ::: :::		
Oy	61 IGRTHHNNLDPEYHHNGSSLAGAAYSLTLTKOGLDKLN-EGLEKERIDAAKCSEFTT	119	:::		
D8	76 IGKRI-EANGLGEAPKRNSSLAGAFAISLTLLKKOLDLKLDELINELAEAKCSEAF	134	:::		
Oy	120 NKIKERTDTGREGCYTDADAFAILLKTNGTYTKGAEBELGKFESVEYLSSAAKEMLANSV	179	:  :::		
D8	135 KRLQDSMAQLGKRINATDAADSKAEALLKTNGTYTKGAETEELEIFKFSVESLSAAKEALNSV	194	:  :::		
Oy	180 KELTSPVAESP 191				
D8	195 KELTSPVAESP 206				

RESULT 11  
 S70278  
 outer surface protein C - Lyme disease spirochete  
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70278  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
 M:J. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in Lyme disease spirochete  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70278  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-191 <Lit>  
 A:Cross-references: EMBL:J142871; NID:9858738; PIDN:AAB37014.1; PID:gi1695229  
 A:Experimental source: strain VS461  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match	36.6%	Score 674.5;	DB 2;	Length 191;
Best Local Similarity	75.8%	Pred. No. 6.5e-25;		
Matches 144;	Conservative 18;	Mismatches 25;	Indels 3;	Gaps 3;
3	CNNSGDKGNTSANSADSEYKGNLTETIKKTIITDSMAVLAVKEVEALLSSIDETAKATG	62		

```

Db      1  CANNNGGADTASTNPDSAGPNTLEISKITTSNAVLAKEVEALLSSIDEL-ARTIG 59
QY      63  KKIHNNGLDTEYNHNSLAGAVAISTLIKOKLDGLKN-EGLKEKIDAAKCESEFTNK 121
Db      60  KKI-ENANGLONENADKNSLAGAVAISTLIKOKLDGLKGLKELKELAEKAKCESEFTNK 118
QY      122  LKEKHTDLGKEGVTDADAKAEALITKTNGTKGAEEELGKLFESVEVLSKAKEMLANSVKE 181
Db      119  LOOSNMDLGKHNATDADSKAEALITKTNGTKGAEEELGKLFESVLSKAKEMLANSVKE 178
QY      182  LTSPVVAESP 191
Db      179  LTSPVVAESP 188

RESULT 12
outer surface protein C - Lyme disease spirochete (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
Accession: I40153
A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in
Infect. Immun. 62, 3213-3221, 1994
A:Reference number: I40153; MUID:94314437
A:Accession: I40153
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <RES>
A:Cross-References: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```

```

Query Match      36.6%; Score 674.5; DB 2; Length 191;
Best Local Similarity 75.5%; Pred. No. 6,5e-25;
Matches 142; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY      5  NSGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKAIGK 64
Db      1  NSGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKAIGK 60
QY      65  IHONNGLDTEYNHNSLAGAVAISTLIKOKLDGLKN-EELKEKIDAAKCESEFTTKK 123
Db      61  IQONGGLAEAGHNGTLTLAGAVITISLIKOKLDGLKSEKLEIEKAKCESEFTTKK 120
QY      124  EKHHTDLGKEGVTDADAKAEALITKTNGTKGAEEELGKLFESVEVLSKAKEMLANSVKE 183
Db      121  GEHAQGLGIEVNTDENKAKALITDADKGAELKELKFAVENLAKAKEMLANSVKE 180
QY      184  SPVVAESP 191
Db      181  SPVVAESP 188

RESULT 13
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
Accession: I40268
A:Title: Homology between Borrelia burgdorferi OspC and members of the family of Borrelia
R:Margolis, N.; Hogan, D.; Cleplak, W.
Gene 143, 105-110, 1994
A:Reference number: I40268; MUID:94259285
A:Accession: I40268
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-References: GB:I25413; NID:9495735; PIDN:AAA22956.1; PID:9495736
C:Genetics:

```

```

A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match      36.6%; Score 673.5; DB 2; Length 211;
Best Local Similarity 74.5%; Pred. No. 8.1e-25;
Matches 143; Conservative 16; Mismatches 32; Indels 1; Gaps 1;

QY      1  MACNNSGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKA 60
Db      17  ISCNNSGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKA 76
QY      61  IGGKIHNNGLDTEYNHNSLAGAVAISTLIKOKLDGLKN-EELKEKIDAAKCESEFT 119
Db      77  IGGKIHNNGLDTEYNHNSLAGAVAISTLIKOKLDGLKN-EELKEKIDAAKCESEFT 136
QY      120  NKLKHTDLGKEGVTDADAKAEALITKTNGTKGAEEELGKLFESVEVLSKAKEMLANSV 179
Db      137  KKLSDNQAELGIEVNTDENKAKALITDADKGAELKELKFAVENLAKAKEMLANSV 196
QY      180  ELTSPVVAESP 191
Db      197  ELTSPVVAESP 208

RESULT 14
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
Accession: S70280
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
Mol. Microbiol. 18, 257-263, 1995
A:Reference number: S70255; MUID:96296448
A:Accession: S70280
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Cross-References: EMBL:I42868; NID:9858735; PIDN:AMB37011.1; PID:91695226
A:Experimental source: strain Z57
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```

```

Query Match      36.5%; Score 672.5; DB 2; Length 193;
Best Local Similarity 77.5%; Pred. No. 8.1e-25;
Matches 148; Conservative 9; Mismatches 31; Indels 3; Gaps 3;

QY      3  CANNNGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKAIG 62
Db      1  CANNNGKDGNTSANSADSVKGPNTLEISKITTSNAVLAKEVEALLSSIDEIAKAIG 59
QY      63  KKIHNNGLDTEYNHNSLAGAVAISTLIKOKLDGLK-NEGKEKIDAAKCESEFTNK 121
Db      60  KKIHNNGLDTEYNHNSLAGAVAISTLIKOKLDGLK-NEGKEKIDAAKCESEFTNK 119
QY      122  LKEKHTDLGKEGVTDADAKAEALITKTNGTKGAEEELGKLFESVEVLSKAKEMLANSV 180
Db      120  LKONHNGLDGEGVTDADAKAEALITKTNGTKGAEEELGKLFESVEVLSKAKEMLANSV 179
QY      181  ELTSPVVAESP 191
Db      180  ELTSPVVAESP 190

RESULT 15
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
Accession: S70277

```

R.; Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
Mol Microbiol 18 357-369 1995

Mol. Microbiol. 18, 257-269, 1995

A; Title: Evidence for lateral transfer and recombination in ospc variation in Lyme disease  
 A; Reference Number: 070355; PMID: 06306448

A; Reference number: S70255; MUID:96296448

A;Accession: S70277

A; Status: nucleic acid sequence not shown

A; molecule type: DNA

A;Residues: 1-194 <LIV>

A; Cross-references: EMBL:L42873; NID:g858740; PIDN:AAB37016.1; PID:g1695231

A; Experimental source: strain SIMON

C; Genetics:

A;Gene: ospC

C: Superfamily: Lyme disease spirochete surface protein C

Query Match	Score	DB 2	Length
36.28	666	194	

Best Local Similarity 74.3%; Pred. NO. 1.6e-24;

Matches 142; Conservative 13; Mismatches 34; Indels 2; Gaps 2;

3 C N N S G K D G - N T S A N S A D E S V K G P N L T E I S K K I T D S N A V L L A V K E V E A L L S S I D E I A K A I 61

Db 1 C N N S G K G D S T N P A D E S A K G P N L T E I S K K I T N S N A F V L A K E V E T L V A S I D E L A T K A I 600

QY 62 GKIHQNGLDTEYNHNGSLAGAYAISTLIKOKLDGKN-EGLEKIDAKKCEFTN 1200

Db 61 GKIKNDGTLLENANHGSSLAGAYAIISNLIKQKLDGLKGLEGLNKEIAEAKNCSEAF TK 120

QY 121 KLKEKHTDLGKEGVTADADAKAAILKTNGTKTKGAELGLFESVEVLSKAKEMLANSVK 180

Db 121 KLKEKHTD LGKENATDEDAKKA I LKTDATKDKGAEELEKLSVSVASLVKAAQEA L TNSVK 180

QV 181 ELTSPVAESP 191

Db 181 ETSPVAFSP 191

Search completed: March 18, 2002, 09:56:46  
Job time: 369 sec

Job time: 369 sec





XX Claim 43; Page 83-84; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 378 AA;

Query Match 100.0%; Score 1859; DB 22; Length 378;  
Best Local Similarity 100.0%; Pred. No. 2, 9e-121;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKRTDSNAVLAVKEVEALLSIDELAANA 60  
1 macnnsygdgnlsansadesvkgpnlteinkrtidsnavllavkeveallsideilaaka 60  
QY 61 IGGKIHQNNGLDTENNHNHNSLAGAVASTLTKOKLDGLKNEGLKEKIDAAKCSFTFN 120  
61 igkkihqngldtennmhngslagayastlikqldglnegllkekidaakkcsftfn 120  
DB 61 igkkihqngldtennmhngslagayastlikqldglnegllkekidaakkcsftfn 120  
QY 121 KIKKHTDLGKEGVTDAKAEALIKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 180  
121 kikekhtdlgkegytdadakeaillkangtktgaeelgklfesvevlskaakemlansvk 180  
DB 121 kikekhtdlgkegytdadakeaillkangtktgaeelgklfesvevlskaakemlansvk 180  
QY 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTESNAVLAVKEV 240  
181 eltsprvaespckpsmvnnsygdgnlsansadesvkgpnlteiskrtesnavlavkev 240  
DB 181 eltsprvaespckpsmvnnsygdgnlsansadesvkgpnlteiskrtesnavlavkev 240  
QY 241 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLSGAVLYLSNLTITKRTKISAKSGELKA 300  
241 etltsidelaikaigkikndvslidneadhnslsgaylylsnltitkrtkisaiksgelka 300  
QY 301 EIEKAKCSEFTAKLKGHTDLGKEGVTDDNAKKAIIKTNDKTKGADELKLFESVKN 360  
301 iekakcseftaklkghetdlgkegytddnakkaiiktndkrtkgaDELKLFESVKN 360  
DB 301 iekakcseftaklkghetdlgkegytddnakkaiiktndkrtkgaDELKLFESVKN 360  
QY 361 LKAKKEMLTNSVKELTS 378  
361 lskakemltnsvkelts 378  
DB 361 lskakemltnsvkelts 378

## RESULT 2

AAB62738  
AAB62738 standard; Protein: 401 AA.

AC AAB62738;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 80.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNIV ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Sehnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI: 2001-050113/06.  
DR N-PSDB: AAF29042.

XX Compositions of ospc polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease.

PS Claim 43; Page 150-151; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 401 AA;

Query Match 98.7%; Score 1834; DB 22; Length 401;  
Best Local Similarity 98.9%; Pred. No. 1, 7e-119;  
Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEINKRTDSNAVLAVKEVEALLSIDELAANA 61  
2 acnnsygdgnlsansadesvkgpnlteinkrtidsnavllavkeveallsideilaaka 61  
DB 25 scnnsygdgnlsansadesvkgpnlteiskrtidsnavllavkeveallsideilaaka 84  
QY 62 IGGKIHQNNGLDTENNHNHNSLAGAVASTLTKOKLDGLKNEGLKEKIDAAKCSFTFN 121  
62 igkkihqngldtennmhngslagayastlikqldglnegllkekidaakkcsftfn 121  
DB 85 gkkihqngldtennmhngslagayastlikqldglnegllkekidaakkcsftfn 144  
QY 122 IKEKHTDLGKEGVTDAKAEALIKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 181  
122 ikekhtdlgkegytdadakeaillkangtktgaeelgklfesvevlskaakemlansvk 181  
DB 145 lkehtdlgkegytdadakeaillkngtktgaeelgklfesvevlskaakemlansvk 204  
QY 182 LTPSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTESNAVLAVKEV 241  
182 ltpspvaespckpsmvnnsygdgnlsansadesvkgpnlteiskrtesnavlavkev 241  
DB 205 ltpspvaespckpsmvnnsygdgnlsansadesvkgpnlteiskrtesnavlavkev 264  
QY 242 TLTSTIDELAKAIGKKIKNDVSLDNEADHNSLSGAVLYLSNLTITKRTKISAKSGELKA 301  
242 tlttsidelaikaigkikndvslidneadhnslsgaylylsnltitkrtkisaiksgelka 301  
DB 265 tlttsidelaikaigkikndvslidneadhnslsgaylylsnltitkrtkisaiksgelka 324  
QY 302 IEKAKCSEFTAKLKGHTDLGKEGVTDDNAKKAIIKTNDKTKGADELKLFESVKN 361  
302 iekakcseftaklkghetdlgkegytddnakkaiiktndkrtkgaDELKLFESVKN 361  
DB 325 iekakcseftaklkghetdlgkegytddnakkaiiktndkrtkgaDELKLFESVKN 384  
QY 362 SKAKKEMLTNSVKELTS 378  
362 skakemltnsvkelts 378  
DB 385 skakemltnsvkelts 401

## RESULT 3

AAB62710  
AAB62710 standard; Protein: 374 AA.

AC AAB62710;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 24.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

```

XX (UNY ) UNIT NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoat G, Dykhulzen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29014.
XX
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 78-79; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
XX chimeric ospc proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 374 AA:
XX
XX Query Match 97.1%; Score 1805.5; DB 22; Length 374;
XX Best Local Similarity 98.1%; Pred. No. 1.5e-117;
XX Matches 370; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
XX
XX 2 ACNNSGKDGNTSANSADSVKGNLTETIRKKTITDSNAVLAVKEVEALLSSIDEIAAKAI 61
XX 1 acnnskgdntsansadesvkgpnltetiskitdsnavllavkeveallssideiaakai 60
XX
XX 62 GKRIHONNGIDTENNHNHNSGLAGAVAI STLKOKLDGLKNEGIEKEIKDAAKKCEFTFNK 121
XX 61 gkrlhngndtelnhngslilagavai stllkqldglneglkexidaakcsetfkn 120
XX
XX 122 LKEKHTDLKEGVTDADAKAAILKANGTKTKGADELKLFESVEVSKAKKEMLAHVKE 181
XX 121 lkehndldgkvgtdadaeaailkngtktkgaeelgklfesvevlskaakemlahvke 180
XX
XX 182 LTPVVAESPKRPSMYNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEVE 241
XX 181 ltpvvaes---pamvnnsgkdntsansadesvkgpnltetiskitdsnavllavkeve 237
XX
XX 242 TLLTSTIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYVLSNLTITKISAIKDSGELKAE 301
XX 238 tlltsidelakalgkikndvslndeadhngslisgavylsnltitkksaiksagelkae 297
XX
XX 302 IEKAKKCEFTAKLKGHEHTDLKEGVTDNNAKKAITLKTNNNDTKGADDELKLFESVKNL 361
XX 298 iekakcseefaklkghehtldgkvgtdnnaakailktndktkgadeleklfesvkn 357
XX
XX 362 SKAAKEMLTNSVKELTS 378
XX 358 skaakemltnsvkelts 374
XX
XX RESULT 4
XX AAB62728
XX ID AAB62728 standard; Protein; 398 AA.
XX
XX AAB62728;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospc protein SEQ ID NO: 60.
XX
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX
XX Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

```

```

XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UNY ) UNIT NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoat G, Dykhulzen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29032.
XX
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 123; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
XX chimeric ospc proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 398 AA:
XX
XX Query Match 97.0%; Score 1802.5; DB 22; Length 398;
XX Best Local Similarity 97.9%; Pred. No. 2.5e-117;
XX Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
XX
XX 2 ACNNSGKDGNTSANSADSVKGNLTETIRKKTITDSNAVLAVKEVEALLSSIDEIAAKAI 61
XX 25 acnnskgdntsansadesvkgpnltetiskitdsnavllavkeveallssideiaakai 84
XX
XX 62 GKRIHONNGIDTENNHNHNSGLAGAVAI STLKOKLDGLKNEGIEKEIKDAAKKCEFTFNK 121
XX 85 gkrlhngndtelnhngslilagavai stllkqldglneglkexidaakcsetfkn 144
XX
XX 122 LKEKHTDLKEGVTDADAKAAILKANGTKTKGADELKLFESVEVSKAKKEMLAHVKE 181
XX 145 lkehndldgkvgtdadaeaailkngtktkgaeelgklfesvevlskaakemlahvke 204
XX
XX 182 LTPVVAESPKRPSMYNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEVE 241
XX 205 ltpvvaes---pamvnnsgkdntsansadesvkgpnltetiskitdsnavllavkeve 261
XX
XX 242 TLLTSTIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYVLSNLTITKISAIKDSGELKAE 301
XX 262 tlltsidelakalgkikndvslndeadhngslisgavylsnltitkksaiksagelkae 321
XX
XX 302 IEKAKKCEFTAKLKGHEHTDLKEGVTDNNAKKAITLKTNNNDTKGADDELKLFESVKNL 361
XX 322 iekakcseefaklkghehtldgkvgtdnnaakailktndktkgadeleklfesvkn 301
XX
XX 362 SKAAKEMLTNSVKELTS 378
XX 382 skaakemltnsvkelts 398
XX
XX RESULT 5
XX AAB62713
XX ID AAB62713 standard; Protein; 377 AA.
XX
XX AAB62713;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospc protein SEQ ID NO: 30.

```

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 XX Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 XX 28-DEC-2000.  
 XX  
 XX 19-JUN-2000; 2000WO-US16915.  
 PF  
 XX 18-JUN-1999; 99US-0140042.  
 PR  
 XX (UUNY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 PI WPI; 2001-050113/06.  
 XX N-PSDB; AAF29017.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -  
 PS Claim 43; Page 86-87; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 377 AA;

Query Match 86.8%; Score 1614.5; DB 22; Length 377;  
 Best Local Similarity 87.8%; Pred. No. 2.6e-104;  
 Matches 331; Conservative 20; Mismatches 25; Indels .1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEINKRTTDSNAVLAVKEVALSSIDEIAAKA 60  
 DB 1 macnsgkdgntsansadesvkgpmlteinkrttdsnavlavkevalssideiaaka 60  
 QY 61 IGGKIHONNGLDPTENNHNHNSILAGAYASTLIKQKLDGKNGELKEKIDAAKCSFTTN 120  
 DB 61 igkkihongldptennhnhsilagayaistlikqkldgkngelkekidaakcsfttn 120  
 QY 121 KLEKHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 180  
 DB 121 klekhtdlgkgyvdadakkalilkanGTKTGaeelgklfesyevlskaakemlansvk 180  
 QY 181 ELTSPVVAESPKKPSMVNNSGDKGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEV 240  
 DB 181 eltspvvaesppkpsmvnnsdgkntsansadesvkgpmlteiskrttesnavlavkev 240  
 QY 241 ELLTSLIDELA-KAIGKRTKNDVSLDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 299  
 DB 241 elltslidelak-kaigkrtkndvslDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 299  
 QY 241 ELLTSLIDELA-KAIGKRTKNDVSLDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 299  
 DB 241 elltslidelak-kaigkrtkndvslDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 299  
 QY 300 AEIEFAKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 359  
 DB 300 aeiefaKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 359  
 QY 301 EKENAKKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 360  
 DB 301 ekenakKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 360  
 QY 360 NLSKAKEMLTNSVKEL 376  
 DB 360 nlskaKEmLTNSVKel 376  
 QY 361 NLSKAKEMLTNSVKEL 377  
 DB 361 nlskaKEmLTNSVKel 377

RESULT 6  
 AAB62739  
 ID AAB62739 standard; Protein; 400 AA.  
 XX

AC AAB62739;  
 XX  
 XX 03-APR-2001 (first entry)  
 DT  
 XX  
 XX  
 DE Borrelia sp chimeric ospC protein SEQ ID NO: 82.  
 XX  
 XX  
 XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 XX Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 XX WO200078966-A1.  
 XX  
 XX 28-DEC-2000.  
 PD  
 XX 19-JUN-2000; 2000WO-US16915.  
 PF  
 XX 18-JUN-1999; 99US-0140042.  
 PR  
 XX (UUNY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 PI WPI; 2001-050113/06.  
 XX N-PSDB; AAF29043.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -  
 PS Claim 43; Page 153; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 400 AA;

Query Match 85.5%; Score 1589.5; DB 22; Length 400;  
 Best Local Similarity 86.7%; Pred. No. 1.5e-102;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKRTTDSNAVLAVKEVALSSIDEIAAKAI 61  
 DB 2 acnnsGKdgntsansadesvkgpmlteinkrttdsnavlavkevalssideiaakai 61  
 QY 62 GKTIHONNGLDPTENNHNHNSILAGAYASTLIKQKLDGKNGELKEKIDAAKCSFTTNK 121  
 DB 62 gktiHONNGldptennhnhsilagayaistlikqkldgkngelkekidaakcsfttnk 121  
 QY 144 IGGKIHONNGLDPTENNHNHNSILAGAYASTLIKQKLDGKNGELKEKIDAAKCSFTTNK 144  
 DB 144 igkkihongldptennhnhsilagayaistlikqkldgkngelkekidaakcsfttnk 144  
 QY 181 ELLTSLIDELA-KAIGKRTKNDVSLDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 181  
 DB 181 elltslidelak-kaigkrtkndvslDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 181  
 QY 145 IGGKIHONNGLDPTENNHNHNSILAGAYASTLIKQKLDGKNGELKEKIDAAKCSFTTNK 144  
 DB 145 igkkihongldptennhnhsilagayaistlikqkldgkngelkekidaakcsfttnk 144  
 QY 182 LSTSPVVAESPKKPSMVNNSGDKGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEV 241  
 DB 182 lstspvvaesppkpsmvnnsdgkntsansadesvkgpmlteiskrttesnavlavkev 241  
 QY 205 LSTSPVVAESPKKPSMVNNSGDKGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEV 241  
 DB 205 lstspvvaesppkpsmvnnsdgkntsansadesvkgpmlteiskrttesnavlavkev 241  
 QY 242 TLLTSLIDELA-KAIGKRTKNDVSLDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 300  
 DB 242 tlltslidelak-kaigkrtkndvslDNEADHNSILSAGAYLISNLITRKISAIKDSGELK 300  
 QY 301 EIEKAKKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 360  
 DB 301 eieKAKKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 360  
 QY 325 KLENAKKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 384  
 DB 325 klenakKCSSEFTAKLGEHTDLGKGVYDADAKKALIKANGTKTGAEELKLFESVEVLSAAKEMLANSVK 384  
 QY 361 LSKAAKEMLTNSVKEL 376  
 DB 361 lskAAKEmLTNSVKel 376  
 QY 385 LSKAAKEMLTNSVKEL 400  
 DB 385 lskAAKEmLTNSVKel 400

```

RESULT 7
AAB62711
ID AAB62711 standard; Protein: 373 AA.
AC AAB62711;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
XX WO200078966-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UNYNY ) UNIV NEW YORK STATE RES FOUND.
PA (BR00-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Selinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PI
XX MPI: 2001-050113/06.
DR N-PSDB: AAF29015.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 81; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 373 AA:
SO

Query Match 84.0%; Score 1561; DB 22; Length 373;
Best Local Similarity 85.9%; Pred. No. 1,3e-100;
Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2

2 ACNNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVALLISTIDEIAKAI 61
1 acnnskgdntsaanaadevkgpnltetlskkitdsnavllavkevallssideiaakai 60
62 GKTIHONGSLDTENNHNHNSGLAGAYAIISLTIKOKIDGLKNBGLKETIDAKKCEFTFK 121
61 gkltihngmldteyhnhngslagayaisltlkqklidglkneglkexidaakkcseftfk 120
122 LKEKHTDLGKEGVTDADAKKAEALIKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVKE 181
121 lkehtldlgkegvtdadakeallkngctklgaeeigklfesvevlskaakemlansvke 180
182 LTPSYVAESPSPKPSMVNNSGKDGNTSANSADSVKGNLTETISKKTITSENNAVYLAKEVE 241
181 ltpsyvaespkspsmvnnskgdgntsansadsvkgnltetiskktitssenavylakeve 237
242 TLTSLIDELA-KAIGKIKITNDVSLNEADHNHNSLISGAVLISNLTKTKISATISGDELKA 300
238 tltslidelatckatgkikigqngylavegngtllagaytiskliltqklidglknesklike 297
301 ETAKAKGCSSEFTTALKGEHTDLGKEGVTDNAKKAAILKTTNNDKTKGADELEKLFESVKN 360

```

Query Match	83.8%	Score 1558;	DB 22;	Length 397;
Best Local Similarity	85.6%	Pred. NO. 2.3e-100;		
Matches 322;	Conservative 23;	Mismatches 27;	Indels 4;	Gaps 2.
<p>2 ACNNGSGKGNISANSADSVKGNPTETNKTTDSNAVLAVKVEYALLSIDETAAKAI 61</p> <p>25 scnmgskdgnnsaadsadvkgnptetlskkltdsnavlavkeveallssideaaakai 84</p> <p>62 GKIKQNNGLDTENNHNGSLLAGAVAIISYLIKOKLDGLKNEGLKEKIDAAKCSFTTNNK 121</p> <p>85 gkktimgngldegyhngslilagayaisltlkqklidglneglnkekldeakcksetfnk 144</p> <p>122 LKEKHTDJKEGSVTDADAEKAILKANGTKTKGAEBELGKLFESVEVLISKAKEMLANSVKE 181</p> <p>145 lkehtldtjgkegvtadadaeakailktngctkkgaeelgklfesvevliskaakemlansvke 204</p> <p>182 LTPSPVAESPKKPSPVNNNSGKDGNNISANSADSVKGNPTETLSKTTESNAVLAVKVE 241</p> <p>205 ltpspvaes--pavvnnsgkdgnsaadsadvkgnptetlskkltesnavlavkele 261</p>				

QY 242 TLFTSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELKA 300  
 DB 262 TLIASIDELATKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELKA 321  
 QY 301 EIERAKKCEEFATKIGKHTDLGEGYTDNNAKAILKTNDKTKGADLEKLFESVKN 360  
 DB 322 KLENAKKCEEFATKIGKHTDLGEGYTDNNAKAILKTNDKTKGADLEKLFESVKN 381  
 QY 361 LSKAKKEMLTNSVKEL 376  
 DB 382 LAKAKEMLTNSVKEL 397

RESULT 9  
 AAB62715  
 ID AAB62715 standard; Protein: 370 AA.

AAB62715:

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 34.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNV) UNIV NEW YORK STATE RES FOUND.

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29019.

PI Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

PS Claim 43; Page 91-92; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 370 AA;

Query Match 82.2%; Score 1528.5; DB 22; Length 370;  
 Best Local Similarity 86.3%; Pred. No. 2.3e-98;  
 Matches 327; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 2 ACNNSGKDNTSANSADSVKGNLTETINKITDSNNAVLLAVEVEALLSSIDEIAAKAI 61  
 DB 1 acnnsykdntsansadesvkgpnlteiskitdsnavllaveveallssidel-akai 59

QY 62 GKTHONNGJDTENNHNHSLAGAYIASTLITKOKLGLK-NEGLKRIIDAKKCEFTFN 120  
 DB 60 gkthnngjdtennhnslagayiaystlittqklslklingseglkkaakcseefst 119

QY 121 KLEKHTDLGEGYTDADAKAELIKANGT-KTGAELEKLFESVSLSKAKKEMLTNSV 179

DB 120 KLDNHAGLIGYVDENAKKAILKANAAGKDKGYVEELEKLSGSLSSIAKAKEMLTNSV 179  
 QY 180 KELTSPVAAESPCKPSVNNNSGKDNTSANSADSVKGNLTETINKITDSNNAVLLAVE 239  
 DB 180 keltspvvhg-----nnsykdntsansadesvkgpnlteiskitdsnavllave 231  
 QY 240 VETLISIDELAIAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELK 299  
 DB 232 veciltidelakaiyikikndvslidneadhnsgslisgayslntitkksaiksagelk 291  
 QY 300 AEIERAKKCEEFATKIGKHTDLGEGYTDNNAKAILKTNDKTKGADLEKLFESVKN 359  
 DB 292 aeieakkkceefatlkigkhtdlggytdnnaakailktndktkgadelkifvesk 351  
 QY 360 NLSKAKKEMLTNSVKELTNS 378  
 DB 352 nlskaakemltinsvkelts 370

RESULT 10  
 AAB62730  
 ID AAB62730 standard; Protein: 394 AA.

AC AAB62730;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 64.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNV) UNIV NEW YORK STATE RES FOUND.

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29034.

PI Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

PS Claim 43; Page 128-129; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 394 AA;

Query Match 82.1%; Score 1525.5; DB 22; Length 394;  
 Best Local Similarity 86.0%; Pred. No. 4e-98;  
 Matches 326; Conservative 11; Mismatches 31; Indels 11; Gaps 4;

QY 2 ACNNSGKDNTSANSADSVKGNLTETINKITDSNNAVLLAVEVEALLSSIDEIAAKAI 61  
 DB 25 acnnsykdntsansadesvkgpnlteiskitdsnavllaveveallssidel-akai 83

QY 62 GKRIHONGDTEENNHNGLAGAVASTLIKOKLDGLK-NEGKKEKIDAAKCEFTFN 120  
DB 84 GKRIKNDGSLDNEANNESSILGAYLSTLCQKLSKNGSEGLKLAaAKCSEFTFC 143  
QY 121 KKEKHTDGGKGVTDADAKKAILKANGT-KTGADELKLFESVEVLSKAEMLANSV 179  
DB 144 KJkdnhaqyglgvtndemkalkknaagkdgyveeleklsgslslekaakemlanv 203  
QY 180 KELTSPVVAESPRKPSMVNNSGKDGNTSANSADSEYKGNLTKITKTESNAVLAKE 239  
DB 204 keltspvvhg-----nsgkdgnltsansadesvknltleiskkltesnavlavke 255  
QY 240 VETLTSTIDELAKATGKTKIKNDVSLDNEADHNGSLTSGAVYLSNLTITKRTSATKDSGELK 259  
DB 256 VETLTSTIDELAKATGKTKIKNDVSLDNEADHNGSLTSGAVYLSNLTITKRTSATKDSGELK 315  
QY 300 AEIERAKKCEEFATAKGEHTDLGEGVYTDNAKRAILKTNDKTKGADLEKLFESVYK 359  
DB 316 AEIEKKEKCEEFATAKGEHTDLGEGVYTDNAKRAILKTNDKTKGADLEKLFESVYK 375  
DB 360 NLSKAKEMLTNSVKELTS 378  
DB 376 NLSKAKEMLTNSVKELTS 394

RESULT 11  
AAB62726  
ID AAB62726 standard; Protein; 384 AA.  
AC AAB62726;  
XX  
DT 03-APR-2001 (first entry)  
DE Borrelia sp chimeric ospC protein SEQ ID NO: 56.  
XX  
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
XX  
OS Chimeric - Borrelia sp.  
XX  
OS Chimeric - Borrelia sp.  
XX  
PN WO200078966-A1.  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000WO-US16915.  
XX  
PR 18-JUN-1999; 99US-0140042.  
XX  
PI (UYNY ) UNIV NEW YORK STATE RES FOUND.  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX  
PI WPI: 2001-050113/06.  
XX N-PSDB: AAF29030.  
XX  
DR Compositions of ospC polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -  
PS Claim 43; Page 117-118; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospC proteins and  
XX chimeric ospC proteins from members of the Borrelia genus. These may be  
XX Borrelia burgdorferi, B. afzeili or B. garinii. These can be used as  
XX vaccines against Borrelia infection, which is spread by ticks and leads  
XX to Lyme disease.  
XX  
XX Sequence 384 AA;  
XX  
XX Query Match 80.5%; Score 1496.5; DB 22; Length 384;

Best Local Similarity 82.0%; Pred. No. 4e-96;  
Matches 309; Conservative 32; Mismatches 33; Indels 3; Gaps 2;  
QY 2 ACNNSGKDGNTSANSADSEYKGNLTKITKTESNAVLAKEVEALLSIDELAKAI 61  
DB 1 acnnsqkdgntlsansadesvknltleiskkltsnavlavveeallssidelaakai 60  
QY 62 GKRIHONGDTEENNHNGLAGAVASTLIKOKLDGLKNEGKKEKIDAAKCEFTFN 121  
DB 61 GKRIHONGDTEENNHNGLAGAVASTLIKOKLDGLKNEGKKEKIDAAKCEFTFN 120  
QY 122 LKEKHTDGGKGVTDADAKKAILKANGT-KTGADELKLFESVEVLSKAEMLANSVKE 181  
DB 121 LKEKHTDGGKGVTDADAKKAILKANGT-KTGADELKLFESVEVLSKAEMLANSVKE 180  
QY 182 LTPSPVVAESPRKPSMVNNSGKDGNTSANSADSEYKGNLTKITKTESNAVLAKEVE 241  
DB 181 LTPSPVVAESPRKPSMVNNSGKDGNTSANSADSEYKGNLTKITKTESNAVLAKEVE 238  
QY 242 TLTSTIDELAKATGKTKIKNDVSLDNEADHNGSLTSGAVYLSNLTITKRTSATKDSGELKAE 301  
DB 239 TLTSTIDELAKATGKTKIKNDVSLDNEADHNGSLTSGAVYLSNLTITKRTSATKDSGELKAE 297  
QY 302 IEKAKKCEEFATAKGEHTDLGEGVYTDNAKRAILKTNDKTKGADLEKLFESVYK 361  
DB 298 IEKAKKCEEFATAKGEHTDLGEGVYTDNAKRAILKTNDKTKGADLEKLFESVYK 357  
QY 362 SKAAKEMLTNSVKELTS 378  
DB 358 SKAAKEMLTNSVKELTS 374

RESULT 12  
AAB62737  
ID AAB62737 standard; Protein; 408 AA.  
AC AAB62737;  
XX  
DT 03-APR-2001 (first entry)  
DE Borrelia sp chimeric ospC protein SEQ ID NO: 78.  
XX  
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
XX  
OS Chimeric - Borrelia sp.  
XX  
OS Chimeric - Borrelia sp.  
XX  
PN WO200078966-A1.  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000WO-US16915.  
XX  
PR 18-JUN-1999; 99US-0140042.  
XX  
PI (UYNY ) UNIV NEW YORK STATE RES FOUND.  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX  
PI WPI: 2001-050113/06.  
XX N-PSDB: AAF29041.  
XX  
DR Compositions of ospC polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -  
PS Claim 43; Page 147-148; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospC proteins and  
XX chimeric ospC proteins from members of the Borrelia genus. These may be  
XX Borrelia burgdorferi, B. afzeili or B. garinii. These can be used as  
XX vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.  
XX  
SQ Sequence 408 AA:

Query Match 80.3%; Score 1493.5; DB 22; Length 408;  
Best Local Similarity 81.7%; Pred. No. 7e-96;  
Matches 308; Conservative 33; Mismatches 33; Indels 3; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEVKGPNLTETIKKTTDSNAVLLAVEALLSSIDEIATAKAI 61  
DB 25 scmsngkdgntsansadesvkgpnlteiskitdsnavllaveallssideiataka 84  
OY 62 GKRIHNNGLDLENHNSGLAGAVASTLIKOKLDGLKNEGKKEIDAKKCEFTFNK 121  
DB 85 gkrihngldleynhngslagayaistlikgldglnneglkeidaakkceftfnk 144  
OY 122 LKEKHTDLGEGVTDADAKKAILKANGTGTGAEELGKLFESVEVLSKAKEMLANSVRE 181  
DB 145 lkehtdlggevtdadakeailktngtktgaeelgklfesvevlskaakemlansvke 204  
OY 182 LTSPVAESPCKPSMWNNSGKDGNTSANSADSEVKGPNLTETISKTITESNAVYLAKEYE 241  
DB 205 ltspvaaespckpfignmsg--gdsastpdesakypnlvtviskkitdsnavllave 262  
OY 242 TLTSIDEIATAKIGKRIKNDVSLDNEADHNSLISGAYLISNLTETKRISAIKDSGELKAE 301  
DB 263 allssideiatakgkrikndvsldeanhnslisgaysliskltsvlnseelk 321  
OY 302 IEKAKKCEEFPAKLKGEHTDLGEGVTDNNAKKAILTNNDKTGADLEKLFESVKUL 361  
DB 322 lkeakccseefpaktlkghtdlggevtdnnakkailltnndktgadelkfevskul 381  
OY 362 SKAKEMLTNSVKELTIS 378  
DB 382 skaakemltnsvkeiltis 398

## RESULT 13

AAB62718  
ID AAB62718 standard; Protein: 368 AA.

AC AAB62718;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 40.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNV ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI: 2001-050113/06.

XX DR N-PSDB: AAF29022.

XX Compositions of ospC polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -

PS Claim 43; Page 99-100; 160pp; English.

XX The present invention provides compositions comprising ospC proteins and  
CC chimeric ospC proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

SQ Sequence 368 AA:

Query Match 80.1%; Score 1489.5; DB 22; Length 368;  
Best Local Similarity 82.5%; Pred. No. 1.2e-95;  
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEVKGPNLTETIKKTTDSNAVLLAVEALLSSIDEIATAKAI 61  
DB 1 acnnskgdgnasansadesvkgpnlteiskitdsnavllavevettllasidelata 60  
OY 62 GKRIHNNGLDLENHNSGLAGAVASTLIKOKLDGLKNEGKKEIDAKKCEFTFNK 121  
DB 61 gkrihngldleynhngslagayaistlikgldglnneglkeidaakkceftfnk 119  
OY 122 LKEKHTDLGEGVTDADAKKAILKANGTGTGAEELGKLFESVEVLSKAKEMLANSVRE 181  
DB 120 lkehtdlggevtdadakeailktngtktgaeelgklfesvevlskaakemlansvke 179  
OY 182 LTSPVAESPCKPSMWNNSGKDGNTSANSADSEVKGPNLTETISKTITESNAVYLAKEYE 241  
DB 180 ltspvahg-----nsgkdgntsansadesvkgpnlteiskitdsnavllave 231  
OY 242 TLTSIDEIATAKIGKRIKNDVSLDNEADHNSLISGAYLISNLTETKRISAIKDSGELKAE 301  
DB 232 tltsideiatakgkrikndvsldeanhnslisgaysliskltsvlnseelk 291  
OY 302 IEKAKKCEEFPAKLKGEHTDLGEGVTDNNAKKAILTNNDKTGADLEKLFESVKUL 361  
DB 292 lkeakccseefpaktlkghtdlggevtdnnakkailltnndktgadelkfevskul 351  
OY 362 SKAKEMLTNSVKELTIS 378  
DB 352 skaakemltnsvkeiltis 368

## RESULT 14

AAB62734  
ID AAB62734 standard; Protein: 392 AA.

AC AAB62734;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 72.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNV ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI: 2001-050113/06.





Tue Mar 19 10:57:44 2002

us-09-596-746a-28.rag

Page 10

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:37 ; Search time 124.19 Seconds  
(without alignments)  
441.679 Million cell updates/sec

Title: US-09-596-746a-24

Perfect score: 1842

Sequence: 1 MACNSGKGNTSANSADAES.....KNLSKAKEMLINSVKELTS 375

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP-archaea:\*
- 2: SP-bacteria:\*
- 3: SP-fungi:\*
- 4: SP-human:\*
- 5: SP-invertebrate:\*
- 6: SP-mammal:\*
- 7: SP-mhc:\*
- 8: SP-organella:\*
- 9: SP-plant:\*
- 10: SP-rodent:\*
- 11: SP-rodent:\*
- 12: SP-virus:\*
- 13: SP-vertebrate:\*
- 14: SP-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	926	50.3	192	2	09S3P3	09S3P3 borrelia bu
2	921	50.0	192	2	09R7B1	09R7B1 borrelia bu
3	910	49.4	191	2	P70818	P70818 borrelia bu
4	908	49.3	193	2	09RR53	09RR53 borrelia bu
5	893	48.5	200	2	09ROR9	09ROR9 borrelia bu
6	885	48.0	182	2	09R7B2	09R7B2 borrelia bu
7	885	48.0	191	2	031120	031120 borrelia bu
8	879	47.7	192	2	09S3P2	09S3P2 borrelia bu
9	879	47.7	210	2	044719	044719 borrelia bu
10	853	46.3	177	2	044999	044999 borrelia bu
11	843.5	45.8	178	2	09REH7	09REH7 borrelia bu
12	829	45.0	175	2	09R7B4	09R7B4 borrelia bu
13	802	43.5	163	2	09R7A9	09R7A9 borrelia bu
14	786	42.7	165	2	09AGB1	09AGB1 borrelia bu
15	773	42.0	159	2	09R7B0	09R7B0 borrelia bu
16	757.5	41.1	211	2	044720	044720 borrelia bu
17	756.5	41.1	193	2	P94237	P94237 borrelia bu
18	707	38.4	193	2	031115	031115 borrelia bu
19	696.5	37.8	193	2	P94234	P94234 borrelia bu

20	683.5	37.1	211	2	049576	049576 borrelia af
21	678.5	36.8	194	2	031122	031122 borrelia bu
22	678.5	36.8	211	2	044977	044977 borrelia bu
23	677	36.8	212	2	09KIM5	09KIM5 borrelia af
24	676.5	36.7	209	2	044671	044671 borrelia af
25	675	36.6	201	2	P96571	P96571 borrelia ja
26	674.5	36.6	191	2	044726	044726 borrelia bu
27	674.5	36.6	191	2	P94245	P94245 borrelia bu
28	674	36.6	201	2	P96573	P96573 borrelia ja
29	673.5	36.6	211	2	044976	044976 borrelia bu
30	672.5	36.5	193	2	P94242	P94242 borrelia bu
31	672.5	36.5	202	2	P96508	P96508 borrelia af
32	672	36.5	203	2	050624	050624 borrelia af
33	672	36.5	212	2	044705	044705 borrelia bu
34	670	36.4	201	2	P96572	P96572 borrelia ja
35	667	36.2	203	2	050619	050619 borrelia bu
36	666	36.2	194	2	P94244	P94244 borrelia bu
37	665	36.1	190	2	P94247	P94247 borrelia bu
38	661	35.9	212	2	044727	044727 borrelia bu
39	660.5	35.9	201	2	09ROR8	09ROR8 borrelia bu
40	658.5	35.7	194	2	031114	031114 borrelia bu
41	657.5	35.7	193	2	P94231	P94231 borrelia bu
42	657.5	35.7	211	2	0926C7	0926C7 borrelia va
43	655.5	35.6	193	2	P94233	P94233 borrelia bu
44	655.5	35.6	209	2	044883	044883 borrelia bu
45	654.5	35.5	211	2	049577	049577 borrelia af

ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
Q9S3P3	Q9S3P3	01-MAY-2000 (TREMELREL. 13, Created)			
AC	Q9S3P3	01-MAY-2000 (TREMELREL. 13, Last sequence update)			
DT	01-MAY-2000 (TREMELREL. 13, Last sequence update)				
DT	01-JUN-2001 (TREMELREL. 17, Last annotation update)				
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
ID	STRAIN-IP2;				
RC	MEDLINE=96296448; PubMed=8709845;				
RX	Livey I., Gibbs C.P., Schuster R., Dornier F.;				
RA	"Evidence for lateral transfer and recombination in OspC variation in				
RT	Lyme disease Borrelia."				
RL	Mol. Microbiol. 18:257-269(1995).				
DR	EMBL; L42887; AAB36995.1; "				
DR	InterPro: IPR001800; Lipoprotein_6.				
DR	Pfam: PF01441; Lipoprotein_6; 1.				
FT	NON_TER	1	1		
FT	NON_TER	192	192		
SQ	SEQUENCE	192 AA;	20287 MW;	118467AC84C7E3D CRC64;	
Query Match					
Best Local Similarity 99.5%; Pred. No. 2.1e-35;					
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	3	CNNSGKGNTSANSADAESVKGPMLEISKRIITDSNVLLAVKEVEALLSSIDEIAKAIG 62			
DB	1	CNNSGKGNTSANSADAESVKGPMLEISKRIITDSNVLLAVKEVEALLSSIDEIAKAIG 60			
QY	63	KTIHNNGLDTEYNNHNSLAGAIASTLIKQKLDGKNEGLKEKIDAKKCEFTTNKL 122			
DB	61	KTIHNNGLDTEYNNHNSLAGAIASTLIKQKLDGKNEGLKEKIDAKKCEFTTNKL 120			
QY	123	KEKHTDLGKEGYTDADAKKKAELTKTKRGAEELKLFESVVLSSKAKKEMLANSVKEL 182			



```

DB 67 ICKIKQNNGLDTEENNHNGLAGAYAISTLIKQKLDGKLNELKREKIDAKKCSFTFN 126
QY 121 KKEKHTDGLKCEGVTDADAKKALIKTNGTKGAEELGKLFESEVYLSRAKAKMLANSVK 180
DB 127 KKEKHTDGLKCEGVTDADAKKALIKTNGTKGAEELGKLFESEVYLSRAKAKMLANSVK 186
QY 181 ELTSPV 187
DB 187 ELTSPV 193

RESULT 5
Q9R0R9 PRELIMINARY: PRT: 200 AA.
ID Q9R0R9;
AC Q9R0R9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).
Plasmid cp26.
Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2E7;
RA MEDLINE=20002545; PubMed=10531219;
Hofmeister E.K., Glas G.E., Childs J.E., Persing D.H.:
"Population dynamics of a naturally occurring heterogeneous mixture of
Borrelia burgdorferi clones";
Infect. Immun. 67:5709-5716(1999).
DR EMBL: AF074464; AAD23911.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW Plasmid.
FT NON_TER 200
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BEB6B CRC64;

Query Match 48.5%; Score 893; DB 2; Length 200;
Best Local Similarity 98.4%; Pred. No. 6.7e-34;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
"borrelia burgdorferi sensu stricto, a bacterial species 'made in the
U.S.A.'?";
Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB1894.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 182
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match 48.0%; Score 885; DB 2; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.4e-33;
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 195 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 254
DB 10 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 69
QY 255 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 314
DB 70 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 129
QY 315 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 374
DB 130 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 189
QY 375 S 375
DB 190 S 190

```

```

JUL 8
33P2
ID 09S3P2 PRELIMINARY; PRT; 192 AA.
AC 09S3P2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=297;
RX MEDLINE=96296448; PubMed=8709845;
RA Livey I., Gibbs C.P., Schuster R., Dörner F.;
RT "Evidence for lateral transfer and recombination in OSCP variation in
RT Lyme disease Borrelia."
RL Mol. Microbiol. 18:257-269(1995).
DR EMBL: LA2893; AAB37001.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT 192
SQ SEQUENCE 192 AA; 20472 MW; 46AC8F93E4DFED6C CRC64;

```

```

Query Match 47.7%; Score 879; DB 2; Length 192;
Best Local Similarity 99.4%; Pred. No. 2.8e-33;
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 195 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 254
DB 10 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 69
QY 255 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 314
DB 70 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 129
QY 315 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 374
DB 130 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 189
QY 375 S 375
DB 190 S 182

```

```

RESULT 9
ID 044719 PRELIMINARY; PRT; 210 AA.
AC 044719;
DT 01-NOV-1996 (TREMblrel. 01, Created)

```

```

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OSCP.
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB19;
RX MEDLINE=95154673; PubMed=7851744;
RA Stevenson B., Barthold S.W.;
RT "Expression and sequence of outer surface protein C among North
RT American isolates of Borrelia burgdorferi."
RL FEMS Microbiol. Lett. 124:367-372(1994).
DR EMBL: U04281; AAC43297.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
SQ SEQUENCE 210 AA; 22526 MW; 2D672991D584E4EE CRC64;

```

```

Query Match 47.7%; Score 879; DB 2; Length 210;
Best Local Similarity 99.4%; Pred. No. 3.1e-33;
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 195 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 254
DB 20 NNSGKGNTSANSADSVKGNLTETISKRITSNVAVLAVKEVELTLTSLDELAKAIGKK 79
QY 255 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 314
DB 80 IKNDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAEIEKAKKCEFTAKLK 139
QY 315 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 374
DB 140 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKKEMLTNSVKELT 199
QY 375 S 375
DB 200 S 200

```

```

RESULT 10
ID 044999 PRELIMINARY; PRT; 177 AA.
AC 044999;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TXGW;
RX MEDLINE=95286481; PubMed=7768799;
RA Hansen K.;
RA Theisen M., Botre W., Mathiesen M.J., Mikkelsen B., Lebech A.M.;
RT "Evolution of the Borrelia burgdorferi outer surface protein OSCP."
RL J. Bacteriol. 177:3036-3044(1995).
DR EMBL: X84783; CAA59254.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT 177
SQ SEQUENCE 177 AA; 18800 MW; 156671B9614E7A2D CRC64;

```

```

Query Match 46.3%; Score 853; DB 2; Length 177;

```



Query Match 43.5%; Score 802; DB 2; Length 163;  
 Best Local Similarity 98.8%; Pred. No. 6,9e-30;  
 Matches 161; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACNNSGDNSTANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSSIDEIAAKAI 61  
 DB 1 SCNNSGDNSTANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSSIDEIAAKAI 60

QY 62 GKRIHONGDTEYNHNSGLAGAYASTLIKOKLDGKNEGLKERTDAKKSEFTTNK 121  
 DB 61 GKRIHONGDTEYNHNSGLAGAYASTLIKOKLDGKNEGLKERTDAKKSEFTTNK 120

QY 122 LKEKHTDLGKEGYVDADAKKALIKTNGTKTKGAELGKLFESV 164  
 DB 121 LKEKHTDLGKEGYVDADAKKALIKTNGTKTKGAELGKLFESV 163

RESULT 14  
 DBI  
 Q9AGBI PRELIMINARY; PRT; 165 AA.  
 AC Q9AGBI;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE OSCP-MODIFIED.  
 GN OSCP.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HB19;  
 RA Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,  
 RA Lawson C.L., Swaminathan S.;  
 RT "Crystal structure of outer surface protein C (OspC) from the Lyme  
 RT disease spirochete, Borrelia burgdorferi.";  
 RL EMBL J. 0:0-0(2001)  
 DR EMBL; AF337548; AA021289.1; -  
 FT VARIANT 61 61 M -> I.  
 FT SEQUENCE 165 AA; 17833 MW; 56E01536D22F61BF CRC64;

Query Match 42.7%; Score 786; DB 2; Length 165;  
 Best Local Similarity 98.2%; Pred. No. 3.7e-29;  
 Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 212 VNGPMLTEISKRTTSNAVLAVKEVEALLSSIDEIAAKIGKKIRNDVSLDNEADHNSL 271  
 DB 1 MKGPMLTEISKRTTSNAVLAVKEVEALLSSIDEIAAKIGKKIRNDVSLDNEADHNSL 60

QY 272 ISGAYILSMILTKKISATDSGELKAELKAKKCESEFTAKLGEHTDLGKEGYTDDNAK 331  
 DB 61 MSGAYILSMILTKKISATDSGELKAELKAKKCESEFTAKLGEHTDLGKEGYTDDNAK 120

QY 332 KALIKTNDKTKGADELKLFESVKNLSKAAKEMLTNSVKELTS 375  
 DB 121 KALIKTNDKTKGADELKLFESVKNLSKAAKEMLTNSVKELTS 164

RESULT 15  
 Q9R7B0 PRELIMINARY; PRT; 159 AA.  
 AC Q9R7B0;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TETS;

RX MEDLINE=97478003; PubMed=9336916;  
 RA Ras N.M., Postic D., Foretz M., Baranton G.;  
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
 RT U.S.A.'?"  
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TETS;  
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;  
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; U91799; AA81896.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1 159  
 FT SEQUENCE 159 AA; 16861 MW; 5285D422C43DC213 CRC64;

Query Match 42.0%; Score 773; DB 2; Length 159;  
 Best Local Similarity 99.4%; Pred. No. 1.4e-28;  
 Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 TSANADESVKGPMLTEISKRTDTSNAVLAVKEVEALLSSIDEIAAKIKTHONGNL 71  
 DB 1 TSANADESVKGPMLTEISKRTDTSNAVLAVKEVEALLSSIDEIAAKIKTHONGNL 60

QY 72 DTEYNHNSGLAGAYASTLIKOKLDGKNEGLKERTDAKKSEFTTNKLERHTDLGK 131  
 DB 61 DTEYNHNSGLAGAYASTLIKOKLDGKNEGLKERTDAKKSEFTTNKLERHTDLGK 120

QY 132 EGYTDADAKKALIKTNGTKTKGAELGKLFESVYLSKA 170  
 DB 121 EGYTDADAKKALIKTNGTKTKGAELGKLFESVYLSKA 159

Search completed: March 18, 2002, 10:10:53  
 Job time: 976 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:42 ; Search time 39.62 Seconds  
(without alignments)  
347.030 Million cell updates/sec

Title: US-09-596-746a-24

Perfect score: 1842  
Sequence: 1 MACNNSGCDGNTSANSADSE.....KNLSRAKEMLTNSVKELTS 375

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues  
Number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database: SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	50.4	210	1	OSCL_BORBU
2	614	33.3	212	1	OSCL_BORBU
3	404.5	22.0	214	1	VM02_BORHE
4	400	21.7	215	1	VM03_BORHE
5	176.5	9.6	1509	1	MYSN_ACAKA
6	170	9.2	1251	1	RBP2_PLAVB
7	169	9.2	1957	1	YD86_SCHPO
8	167	9.1	483	1	M6_STRPY
9	167	9.1	564	1	M12_STRPY
10	163.5	8.9	1780	1	USOL_YEAST
11	162.5	8.8	3210	1	CENF_HUMAN
12	159	8.6	2116	1	MYSS2_DICDI
13	157.5	8.6	1786	1	LM01_MOUSE
14	157	8.5	998	1	SCAA_RICAK
15	157	8.5	1938	1	MYSS_ABOIR
16	155	8.4	705	1	YNPO_CAEEL
17	154	8.4	1966	1	MYSS_CAEEL
18	153	8.3	1427	1	REST_HUMAN
19	152	8.3	492	1	M5_STRPY
20	150.5	8.2	775	1	YHGE_BACSU
21	148.5	8.1	1713	1	LM03_HUMAN
22	148	8.0	1433	1	REST_CHICK
23	147	8.0	1164	1	BAG_STRAG
24	146.5	8.0	875	1	ZIP1_YEAST
25	146	7.9	1786	1	LM01_HUMAN
26	146	7.9	2349	1	TPR_HUMAN
27	145.5	7.9	1500	1	SSP5_STRGN
28	145.5	7.9	1961	1	MYSS_RAT
29	145	7.9	444	1	PUR2_MERJA
30	145	7.9	962	1	VDP_HUMAN
31	145	7.9	2869	1	RBP1_PLAVB
32	145	7.9	3110	1	LM02_HUMAN
33	144	7.8	1189	1	SCIL_CHICK

34	144	7.8	2017	1	MYSN_DROME	099323 drosophila
35	142.5	7.7	679	1	YK9_YEAST	P34237 saccharomyc
36	142.5	7.7	3106	1	LM02_MOUSE	060675 mus musculu
37	142	7.7	1084	1	MYSS_RABIT	P02562 oryctolagus
38	142	7.7	1935	1	MYSS_CYPCA	090339 cypripinus ca
39	141.5	7.7	961	1	VDP_BOVIN	P41541 bos taurus
40	141.5	7.7	1013	1	SCAA_RICRH	094181 rickettsia
41	141	7.7	1186	1	SMC_BACSU	P51834 bacillus su
42	140.5	7.6	1067	1	EG52_XENLA	091783 xenopus lae
43	140	7.6	1312	1	RA50_YEAST	P12753 saccharomyc
44	139.5	7.6	1060	1	EG51_XENLA	P28025 xenopus lae
45	139.5	7.6	1526	1	MYSS2_SCHPO	G9516 schizosacch

## ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD	PRT	210 AA
AC	007337				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB19.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OG	Plasmid 1p54.				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID:139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-93239332; PubMed-8098841;				
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,				
RA	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (ospc) and the flagellin of Borrelia burgdorferi.";				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-93239332; PubMed-8478108;				
RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant				
RT	major outer surface protein of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-94041630; PubMed-8225587;				
RA	Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.";				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-98065943; PubMed-9403685;				
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Goeysne J.D., Weidman J.,				

RA Uterback T., Matthey L., McDonald L., Attiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "genomic sequence of a Lyme disease spirochete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X69596; CAA49306.1; -  
 CC EMBL: 001894; AAA16058.1; -  
 CC EMBL: D49487; BAA08457.1; -  
 CC EMBL: AE00792; AAC66329.1; -  
 CC TIGR: BBH19; -  
 CC InterPro: IPR001800; Lipoprotein\_6.  
 CC Pfam: PF01441; Lipoprotein\_6; 1.  
 CC ProDom: PD001149; Lipoprotein\_6; 1.  
 CC PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 CC Complete proteome.  
 CC KW SIGNAL 1 18 BY SIMILARITY  
 CC CHAIN 19 210 OUTER SURFACE PROTEIN C.  
 CC LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 CC SEQUENCE 210 AA; 22340 MW; 7A4FC97BF9177BF CRC64;  
 SQ  
 Query Match 50.4%; Score 928; DB 1; Length 210;  
 Best Local Similarity 98.4%; Pred. No. 1.3e-35;  
 Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 MACNNSGKGDTNSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 60  
 Db :::  
 Oy 17 ISCNNSGKGDTNSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 76  
 Db :::  
 Oy 61 IGKRIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGKNEGLKEDIKIDAKKCSFTFN 120  
 Db :::  
 Oy 77 IGKRIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGKNEGLKEDIKIDAKKCSFTFN 136  
 Db :::  
 Oy 121 KLEKHHDLGREGVTDADAKAELIKTNGTKGAELIGKLFESVEVLSKAEMLANSVK 180  
 Db :::  
 Oy 137 KLEKHHDLGREGVTDADAKAELIKTNGTKGAELIGKLFESVEVLSKAEMLANSVK 196  
 Oy 181 ELTSPVVAESP 191  
 Db :::  
 Db 197 ELTSPVVAESP 207  
 RESULT 2  
 OSC2\_BORBU STANDARD; PRT; 212 AA.  
 AC 008137;  
 DT 15-DEC-1998 (rel. 37, Created)  
 DT 15-DEC-1998 (rel. 37, Last sequence update)  
 DT 15-DEC-1998 (rel. 37, Last annotation update)  
 DE OUTER SURFACE PROTEIN C PRECURSOR (PC).  
 GN OSpC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=PKO;  
 RX MEDLINE=92219955; PubMed=1560779;  
 RA Fuchs R., Jauris S., Lottspeich F., Praeger-Mursic V., Wilske B.,

RA Soutschek E.;  
 RT "Molecular analysis and expression of a Borrelia burgdorferi gene  
 RT encoding a 22 kDa protein (pc) in Escherichia coli.";  
 RL Mol. Microbiol. 6:503-509(1992).  
 CC [2]  
 CC SEQUENCE OF 1-205 FROM N.A.  
 CC STRAIN=DK26;  
 CC MEDLINE=94075528; PubMed=8253951;  
 CC Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
 RT "Polymorphism in ospC gene of Borrelia burgdorferi and  
 RT immunoreactivity of ospC protein: implications for taxonomy and for  
 RT use of ospC protein as a diagnostic antigen.";  
 RL J. Clin. Microbiol. 31:2570-2576(1993).  
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X62162; CAA44093.1; -  
 CC EMBL: X73624; CAA52003.1; -  
 CC InterPro: IPR001800; Lipoprotein\_6.  
 CC Pfam: PF01441; Lipoprotein\_6; 1.  
 CC ProDom: PD001149; Lipoprotein\_6; 1.  
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.  
 CC KW SIGNAL 1 18 BY SIMILARITY  
 CC CHAIN 19 212 OUTER SURFACE PROTEIN C.  
 CC LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 CC SEQUENCE 212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;  
 SQ  
 Query Match 33.3%; Score 614; DB 1; Length 212;  
 Best Local Similarity 69.9%; Pred. No. 1.4e-21;  
 Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2;  
 Oy 1 MACNNSGKGDTNSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 59  
 Db :::  
 Oy 17 ISCNNSGKGDSASINPADESAGPNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 76  
 Db :::  
 Oy 60 AIGKRIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGKN-EGLEKEDIKIDAKKCSFTF 118  
 Db :::  
 Oy 77 AIGKRIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGKN-EGLEKEDIKIDAKKCSFTF 136  
 Db :::  
 Oy 119 TNKLEKHHDLGREGVTDADAKAELIKTNGTKGAELIGKLFESVEVLSKAEMLANSVK 178  
 Db :::  
 Db 137 TNKLEKHHDLGREGVTDADAKAELIKTNGTKGAELIGKLFESVEVLSKAEMLANSVK 196  
 Oy 179 VKELTSPVVAESP 191  
 Db :::  
 Db 197 VKELTSPVVAESP 209  
 RESULT 3  
 VM24\_BORHE STANDARD; PRT; 214 AA.  
 AC P32778;  
 DT 01-OCT-1993 (rel. 27, Created)  
 DT 01-OCT-1993 (rel. 27, Last sequence update)  
 DT 01-FEB-1994 (rel. 28, Last annotation update)  
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.  
 GN VMP24.  
 OS Borrelia hermsli.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=140;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-SSP, HSI SEROTYPE 24;  
 RX MEDLINE-93133110; PubMed-1484486;  
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids  
 are highly polymorphic."  
 RL Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L04786; AAA22964.1; -  
 CC Interpro: IPR001800; Lipoprotein\_6.  
 CC Pfam: PF01441; Lipoprotein\_6; 1.  
 CC Prodom: PD001149; Lipoprotein\_6; 1.  
 CC DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 CC KM Outer membrane; Lipoprotein; Signal; Plasmid.  
 CC FT SIGNAL 1 18 PROBABLE.  
 CC FT CHAIN 1 214 VARIABLE MAJOR OUTER MEMBRANE  
 CC FT LIPOPROTEIN 24.  
 CC FT SO LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 CC SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;  
 CC -----  
 CC Query Match 22.0%; Score 404.5; DB 1; Length 214;  
 CC Best Local Similarity 45.4%; Pred. No. 3.1e-12;  
 CC Matches 93; Conservative 36; Mismatches 59; Indels 17; Gaps 5;  
 CC -----  
 CC 1 MACNSGKNGTNSANSADSVKGP----NTEISKTKTDSNAVLAVKEVALLSSIDEI 56  
 CC Db 17 MSCNNGSGPE-----LKSDEVAKSDGTVDLAKVSKRIKESAPFAVKEVTLVKSVDL 71  
 CC QY 57 AAKAIGKKI-HONNGDTEVNHNSLLAGAVAI STLKOKL-----DGLNKGLEKIDA 110  
 CC Db 72 -AAAIKRIKINDGDLDTENGONGSLAGVHVSVAVKIKGALETTSGISINE-LKTITE 129  
 CC QY 111 AKKCSFTFNKLEKHTDLGEGVTDADAKKALIKTNGTKGAELKLFESYEVLKA 170  
 CC Db 130 VKSKAEAFLEKLDKGTHTLGKDDDTKKAIKKNSDKTGASELEALNTAVDALLKA 189  
 CC 171 AKEMLANSVKELTSPVVAESPAMVN 195  
 CC Db 190 AGEVEAIAIKELTAPVKAERPSQNN 214  
 CC -----  
 CC RESULT 4  
 CC VMO3\_BORHE STANDARD; PRT; 215 AA.  
 CC ID VMO3\_BORHE  
 CC AC Q02448;  
 CC DT 01-JUL-1993 (Rel. 26, Created)  
 CC DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 CC DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 CC DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.  
 CC GN VMP3.  
 CC OS Borrelia hermsli.  
 CC OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 CC OX NCBI\_Taxid=140;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN-SSP, HSI SEROTYPE 3;  
 CC RX MEDLINE-93133110; PubMed-1484486;  
 CC RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 CC RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

RT are highly polymorphic."  
 RL Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP24.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L04789; AAA22967.1; -  
 CC Interpro: IPR001800; Lipoprotein\_6.  
 CC Pfam: PF01441; Lipoprotein\_6; 1.  
 CC Prodom: PD001149; Lipoprotein\_6; 1.  
 CC DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 CC KM Outer membrane; Lipoprotein; Signal; Plasmid.  
 CC FT SIGNAL 1 18 PROBABLE.  
 CC FT CHAIN 1 215 VARIABLE MAJOR OUTER MEMBRANE  
 CC FT LIPOPROTEIN 3.  
 CC FT SO LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 CC SEQUENCE 215 AA; 23139 MW; 684C74D35F87C771 CRC64;  
 CC -----  
 CC Query Match 21.7%; Score 400; DB 1; Length 215;  
 CC Best Local Similarity 45.9%; Pred. No. 5e-12;  
 CC Matches 94; Conservative 34; Mismatches 61; Indels 16; Gaps 5;  
 CC -----  
 CC 1 MACNSGKNGTNSANSADSVKGP----NTEISKTKTDSNAVLAVKEVALLSSIDEI 56  
 CC Db 17 MSCNNGSGPE-----LKSDEVAKSDGTVDLAKVSKRIKIDSDFAVKEVHTLVKSIDEI 71  
 CC QY 57 AAKAIGKKI-HONNGDTEVNHNSLLAGAVAI STLKOKL-----NEGLEKIDA 110  
 CC Db 72 -AAAIKRIKINDGDLDTENGONGSLAGVHVSVAVKIKGALETTSGISINE-LKTITE 129  
 CC QY 111 AKKCSFTFNKLEKHTDLGEGVTDADAKKALIKTNGTKGAELKLFESYEVLKA 170  
 CC Db 131 VKKSEAFVQVSKKHHDLAKEGVTDAAHAKSALLVTGTOKKGAELIKLNTAIDELKA 190  
 CC 171 AKEMLANSVKELTSPVVAESPAMVN 195  
 CC Db 191 ANDAVETVIRKELTAPVKAERPSQNN 215  
 CC -----  
 CC RESULT 5  
 CC MYSN\_ACACA STANDARD; PRT; 1509 AA.  
 CC ID MYSN\_ACACA  
 CC AC P05659;  
 CC DT 01-NOV-1988 (Rel. 09, Created)  
 CC DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 CC DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 CC DE MYOSIN II HEAVY CHAIN, NON MUSCLE.  
 CC OS Acanthamoeba castellanii (Amoeba).  
 CC OX Eukaryota; Acanthamoebidae; Acanthamoeba.  
 CC OX NCBI\_Taxid=5755;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE-87308395; PubMed-3040773;  
 CC RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;  
 CC RT "Complete nucleotide sequence and deduced polypeptide sequence of a  
 CC RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a  
 CC RT hinge in the rodlike tail."  
 CC RL J. Cell Biol. 105:913-923(1987).  
 CC CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MLC-2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS  
 CC INTERPRETED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE  
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.  
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING  
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: Y00624; CAA68663.1; .  
 DR PIR: A27224; A27224.  
 DR HSSP: P08799; IAMD.  
 DR InterPro: IPR000048; IQ.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00612; myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR Myosin; coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;  
 KW Methylation; Alkylation; Phosphorylation; Multigene family.  
 FT DOMAIN 1 789  
 FT DOMAIN 790 819  
 FT DOMAIN 848 1509  
 FT DOMAIN 848 1226  
 FT DOMAIN 1227 1252  
 FT DOMAIN 1253 1509  
 FT DOMAIN 1253 1482  
 FT DOMAIN 1483 1509  
 FT NP\_BIND 182 189  
 FT DOMAIN 660 682  
 FT DOMAIN 766 780  
 FT MOD\_RES 133 133  
 FT MOD\_RES 700 700  
 FT MOD\_RES 1489 1489  
 FT MOD\_RES 1494 1494  
 FT MOD\_RES 1499 1499  
 FT MOD\_RES 171201 MW; 2CE49BE51173D17E CRC64;  
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;  
 Query Match 9.6%; Score 176.5; DB 1; Length 1509;  
 Best Local Similarity 24.8%; Pred. No. 0.41;  
 Matches 103; Conservative 68; Mismatches 137; Indels 107; Gaps 21;  
 OY 28 ELSKRTDSNAVLAVKEVEA-----LSSIDE-----IAKAIQKTHQNGDTEYNH 77  
 DB 934 ELQETSSANDILEQKRLLEAKGELKASLEERENKRALOEAKTYESENENLODYED 993  
 OY 76 NSLLAGYAISTLIKOKLDLKNGLKEKIDA--AKKSEFTTNKLEKHTDLGKEGV 134  
 DB 994 EA-----AAHSDLKKEEEDLSRE-LRETKDALDAENISETLRSLK-KNTERGADV 1043  
 OY 135 TDA-DAKELIKTNGTGTGAEELGKLFESYEVLSKAEMLANSVK----- 181  
 DB 1044 RNEEDDTATLQLEKTKKSLFEELAGTRQOLEE-EKSGKAASSKAKGOLGOOLEPARSE 1102  
 OY 182 ---LTSPPVAESPAMVNSGKDN-----TSANSADSVKG--DNLTEISKIT 225  
 DB 1103 VDSIKSKLSAERSL-KTAKDNRDLEDEQLDEERTVRAVNDOKKALEKTELEDOVT 1160

OY 226 ---ESNAVLAVKEVELTILSIDEL-----AKAIQKIKNDVSLDNEADHNSGL 271  
 DB 1161 ALDQGRKA---AAQAKTLTQYDETRRLEEAASAAREKERN--ALDEVAQ----- 1210  
 OY 272 ISGAVYLSNLITKIKSAIKDSG-----ELKAEIERAKK---CSEFTKLKGE 316  
 DB 1211 -----LTADLDAERDSGAQORRKLNTRISELSQSLFNAPKTCGASSEVVRLEGE 1260  
 OY 317 HTDLGKEGVDDAKKAILKTNDKTKGADLEKLFESYKNSKAKEMLINSVK 371  
 DB 1261 LELEBELLTQAPRAAKR-NIDKAN--LELELRQEAADARNDKLVKDNKR 1312  
 RESULT 6  
 ID RBP2\_PLAVB STANDARD; PRT; 1251 AA.  
 AC 000799;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).  
 GN RBP2.  
 OS Plasmodium vivax (strain Belen).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 OX NCBI\_TaxID=31273;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92315338; PubMed=1617731;  
 RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
 RT "A reticulocyte-binding protein complex of Plasmodium vivax  
 RT merozoites".  
 RT Cell 69:1213-1226(1992).  
 CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
 CC HUMAN RETICULOCYTE CELLS.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M88098; AAA29744.1; .  
 KW Malaria; Receptor; Membrane.  
 FT NON\_TER 1  
 FT NON\_TER 1251  
 FT NON\_TER 1251  
 FT NON\_TER 1251  
 SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;  
 Query Match 9.2%; Score 170; DB 1; Length 1251;  
 Best Local Similarity 22.9%; Pred. No. 0.65;  
 Matches 93; Conservative 50; Mismatches 139; Indels 124; Gaps 17;  
 OY 22 KGNLTETSKRTDSNAVLAVKEVEALLSSIDEIYAKA--IGKKTHQNGDTEYNH 79  
 DB 774 KNTNEDVHKNIDAVKVALEI-----LAHDEIDTRKQKSSKLIEMNQI----- 819  
 OY 80 SILAGYAISTLIKOKLDLKNGLKEKIDA--AKKSEFTTNKLEKHTDLGKEGV 136  
 DB 820 -----YLRVVLINQ-----YKKNISSTKSEEVAVYKGVNSKSHSELSTKCS 864  
 OY 137 ADAKEAIL-----KTNQTKTKGAELGKLFESYEVLSKAEMLANSV 179  
 DB 865 KSYNDIALKQTELQNLNSPTQKNTNNSKLEKTKTFES-----LKNAL 913  
 OY 180 KLTSPVAESPAMVNSGKDNSTANSADSVKGNLTETSKRTDSNAVLAVKEVEY 239  
 DB 914 KTLGEEVNA-----LKASDNEHGVQKSEPV-NPALSEIEKEBT-----DIDS 956  
 OY 240 LITSIDELAK-----AIGKTKNDVSLDNEADHNSGLISGAYL----- 277

```

Db 957 LMTALDELLKGGRTCEVSRKILKDTVTKEISDTELIINTIEKNV---AYLAIYIKNV 1012
Qy 278 ---ISLITL--KTSIAIKDSGELKAEIERAKKCSSEPTAKLKEHTDLGEGVTDNAKK 332
Db 1013 EDTVOOVLTNLNEHFNTRKQVSNHEPTNFKSNKSEELTKAVTDSKTIISK-----LKG 1065
Qy 333 AILKTN-----NDKTGKADELEKLFESVKNLSAAKEM--LTNSVK 371
Db 1066 VIEIVNEMTNTIESSAKEIEALYNELKNKKTSLNEITQTSNEVK 1111

RESULT 7
YD86_SCHPO STANDARD: PRT: 1957 AA.
ID YD86_SCHPO
AC Q10411:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
OC Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 270690; CAA94624.1;
DR Hypothetical protein.
KW SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;
SQ

```

```

Db 690 NIOTLEKEDLKKSEBALRFSLEAKNLREVIDNLKGKHETLEAQNDHSSLSDAKKNTNAI 749
Qy 335 LKTNNDKTKGADELEKLFESVKNL---SKAAKEMLTNSVELTS 375
Db 750 L--SSLETKSESDYKRLTANVETLTQDSKAMKQSFSTLVNSYGS 791

RESULT 8
M6_STRPY STANDARD: PRT: 483 AA.
ID M6_STRPY
AC P08089;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 6 PRECURSOR.
GN EMM6.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacilllus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8611835; PubMed=3511046;
RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT "Complete nucleotide sequence of type 6 M protein of the group A
RT Streptococcus. Repetitive structure and membrane anchor.";
RL J. Biol. Chem. 261:1677-1686(1986).
RN [2]
RP SEQUENCE OF 43-122 FROM N.A.
RX MEDLINE=8516224; PubMed=3885219;
RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT "Relationship of M protein genes in group A streptococci.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC PHAGOCYTOSIS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: TO OTHER M PROTEINS.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M1138; AAA26920.1;
DR PIR: A26297;
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PRINTS: PR00015; GP0SANCHOR.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW Transmembrane; Coiled coil; Signal.
FT SIGNAL 1 42
FT CHAIN 43 483
FT DOMAIN 43 457
FT TRANSMEM 458 477
FT DOMAIN 478 483
FT DOMAIN 478 483
FT DOMAIN 69 138
FT DOMAIN 157 269
FT DOMAIN 279 347
FT DOMAIN 348 411
FT DOMAIN 412 446
FT DOMAIN 449 454

```

FT PROTEINS.  
SQ SEQUENCE 483 AA: 53472 MW: 6687FF28DB53AA48 CRC64:

Query Match 9.1%; Score 167; DB 1; Length 483;  
Best Local Similarity 23.3%; Pred. No. 0.31;  
Matches 91; Conservative 57; Mismatches 157; Indels 86; Gaps 14;

QY 5 NSGCGNTSANSADSVKGNLTETISKRTDSNAVLLAVVEVALLSSIDEIAKAIGK 64  
DB 73 NNDKLTENNLDON---KNLTENNLDONKNLTENN---NL 112  
QY 65 IHONNGDTEYNHNGSLAGAVASTLIKQKLDGKNEGKEKIDAKKSEFTTNK 124  
DB 113 TDONKNLTET---NKEIAEENRLT---EKKGITKLSE---EBAKERE 157  
QY 125 KHPDGLK-BEVTADAKAELKNGTKTKAEELGLFESVE---VLSKAKE---LAN 177  
DB 158 NKEAIGTKLTDETVKDKIAKQESK---ETIGTLKLTDETVKDKIAKQESKETIG 213  
QY 178 SVKELTSPVVAESPAMVNNSGKDNNTSANSADSVKGNLTETISKRTESNAVLLAVKEV 237  
DB 214 TLKKTIDETVKKDKIAKQESKETIGTKLTDETVK---DKTARE-QSKODIGALKOE 268  
QY 238 ETLTSLIDEIAKAIGKINDVSLDNEADHNGSLISGAVLISMLTKRTSAIKSGELKA 297  
DB 269 LAKKDEGNKVSSEASRKGRLDLDASREA---KQVEKDLANLTA 309  
QY 298 EIRKAKKSEFTTNKLTGEGVTDNNAKKAIIKLTNDKTKGADELEKLFESVKN 357  
DB 310 ELDKVKEKQISDASRQGLRDLASREAKOVERKA-LEBANSKIALALEKLEESK 368  
QY 358 L-----SKAKEMLTNSVKEL 373  
DB 369 LTEKERAELOAKLEBAKALKLEQAKQAEEL 399

RESULT 9  
M12\_STRPY STANDARD; PRT; 564 AA.  
AC P19401;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).  
GN EMM12.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
STREPTOCOCCUS.  
OX NCBI\_TaxID=1314;

RP SEQUENCE FROM N.A.  
RC STRAIN-CS24 / Serotype M12;  
RA MEDLINE=88058777; PubMed=2445730;  
RX Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;  
RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream  
sequences";  
RL J. Bacteriol. 169:5633-5640(1987).  
CC -I- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES  
OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF  
THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO  
PHAGOCYTOSIS.  
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.  
CC -I- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS  
IN THE REGION OF THE MEMBRANE ANCHOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb.ch/announce/>

CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch).

DR EMBL: M1269; AA88573.1; -  
DR PIR: A60115; A60115.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR003345; M\_repeat.  
DR Pfam: PF00746; Gram\_pos\_anchor. 1.  
DR Pfam: PF02370; M: 9.  
DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
KW Transmembrane; Coiled coil; Signal.  
FT SIGNAL 1 41  
FT CHAIN 42 564 M PROTEIN, SEROTYPE 12.  
FT DOMAIN 42 550 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 551 >564 MEMBRANE ANCHOR.  
FT DOMAIN 44 505 COILED COIL (POTENTIAL).  
FT DOMAIN 505 541 GLY/PRO-RICH (CELL WALL-SPANNING).  
FT DOMAIN 542 547 CONSERVED IN GRAM-POSITIVE COCCI SURFACE  
FT PROTEINS.  
FT NON\_TER 564 564  
SQ SEQUENCE 564 AA: 62904 MW: 5F1549DACA77B46 CRC64;

Query Match 9.1%; Score 167; DB 1; Length 564;  
Best Local Similarity 22.6%; Pred. No. 0.36; Mismatches 161; Indels 116; Gaps 17;

Matches 97; Conservative 55; Mismatches 161; Indels 116; Gaps 17;  
QY 10 GNTSANSADSVKGNLTETISKRTDSNAVLLAV-KEVEALLSSIDEIAKAIGKIHON 68  
DB 115 GRLGIDNMDLKA---ITELKSVKENDVLSQIKKELEAEKIDQ-----FGREYHAA 165  
QY 69 NGLDTEYNHNGSLAGAVASTLIKQKLDGKNEGKEKIDAKKSEFTTNKLEKHTD 128  
DB 166 DIL---NKKQELAKENVIS---KLNG-ELQPLKQVD---ETDNNLOE--- 205  
QY 129 LKKEGVTDADAKAELKNGTKTKGAELGLFESVEVLSRAA---KEMIANSVKELTSPV 186  
DB 206 --KQKVLSEQQLAVTKENAKKDFELALGHQLDKEVNAKIALESESLADAKKDFELAA 263  
QY 187 V---AESPMVNNSGKNG-----NTSANSADSVKGP-----NL 218  
DB 264 LGHOHANEYQAKLEKQIKOLEEKOILDASRKGTARDLEAVRAKKAETAEINLNLK 323  
QY 219 EISKRTESNAVLLAVK-----EVELTTSIDE---LAKAIGKINDV 259  
DB 324 AELAKVTEQKQILDSRGTARDLEAVRKKQVLEAKQLEQNNKISEASRKGRLD 383  
QY 260 SLDNEADHNGSLISGAVLISMLTKRTSAIKDSGELAEIEKAKKSEFTTNKLEKHTD 319  
DB 384 DTSREA---KQVEKDLANLTAELDKVKEEKQISDASRQGLRD 424  
QY 320 LKKEGVTDNNAKKAIIKLTNDKTKGADELEKLFESVKNL-----SKAKE 364  
DB 425 LDASREAKOVERKA-LEBANSKIALALEKLNDEESKRLTEKERAELOAKLEBAKALKE 483  
QY 365 MLTNSVKEL 373  
DB 484 QIAKQAEEL 492

RESULT 10  
USO1\_YEAST STANDARD; PRT; 1790 AA.  
AC P25386;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.  
GN USO1 OR INT1 OR YDI058W.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-X2180-1A;  
 RX MEDLINE-91183402; PubMed-2010462;  
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
 RA Yamasaki M.;  
 RT "A cytoskeleton-related gene, usol, is required for intracellular  
 RT protein transport in *Saccharomyces cerevisiae*.";  
 RL J. Cell Biol. 113:245-260(1991).  
 RN [2]  
 RP SEQUENCE OF 782-1790 FROM N.A.  
 RA Hosetier M.K., Herman D.J., Bendel C.M., McCellan M., Tao N.,  
 RA Kendrick K.E.;  
 RL Submitted (Feb-1993) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;  
 RL Submitted (May-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL: X54378; CA38253.1; -  
 DR EMBL: L03188; AAB00143.1; -  
 DR EMBL: U53668; AAB6659.1; -  
 DR PIR: A38455; A38455.  
 DR HSSP: P80220; 1D1P.  
 DR SGD: S0002216; USOL.  
 DR Interpro: IPR002017; Spectrin.  
 KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
 FT DOMAIN 1 724 GLOBULAR HEAD.  
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).  
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).  
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.  
 FT DOMAIN 1172 1785 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 847 847 G->E (IN REF. 2).  
 FT CONFLICT 924 924 E->K (IN REF. 2).  
 FT CONFLICT 1253 1253 V->I (IN REF. 2).  
 FT CONFLICT 1319 1319 I->V (IN REF. 2).  
 FT CONFLICT 1461 1461 N->S (IN REF. 2).  
 FT CONFLICT 1581 1581 G->S (IN REF. 2).  
 FT CONFLICT 1600 1600 I->V (IN REF. 2).  
 FT CONFLICT 1661 1661 R->S (IN REF. 2).  
 FT CONFLICT 1772 1772 D->DEEDDEE (IN REF. 2).  
 SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 8.9%; Score 163.5; DB 1; Length 1790;  
 Best Local Similarity 21.8%; Pred. No. 1.9;  
 Matches 102; Conservative 73; Mismatches 184; Indels 109; Gaps 17;

OY 3 CUNSGKDG-----MTSANSADSVGPNLTETSKITPSNVLAVKVEVALSS 52  
 DB 929 CUNLSKSKHISLVEYKSRFSHNLV--AKLTSLKSLANNYKDMQA--ENESLTKA 984  
 OY 53 IDEIAAKAIGKIKHONGLD--TEYNHNGSLLAGAY-----AISTL-----IKOR 95  
 DB 985 VEESKNESSIOLSNLQNKIDMSQEKENQIERGSIETKNIETQLKRTISDLEQTKREELISK 1044

OY 96 LDGLKNGC-----LKEIDAAKKSEFTFNKLK-----KHTDLKREGVTD 136  
 DB 1045 SDSSKDEYESQISLKEKLELTATTANDENVKNSLSELTREBELAEALAKNLKNELET 1104  
 OY 137 ADAREKALK-----TNGTKTK-----GAEELCKLESEVFLSKAA 171  
 DB 1105 LEISEKALKKVEKNEHLEKKEKIQLEKEATREYQOQLNSLANLESLEKHEHDLAAQKKY 1164  
 OY 172 KEMLAANSVKELTSPV-----VAESPANVNSKGDGNTSANSADSVGPNLTETSKRITTE 226  
 DB 1165 EEOIANKEROYNEEISOLNDELITSTOENESIKKKNDLEGEYKAMKSTSEOSNLKSE 1224  
 OY 227 SNAVVLAVKREV-----ETLNSIDELAKAIGKIKIKNDVSLDNEADHNSLSGAVLISNLI 282  
 DB 1225 IDALNTQIKELKKNKNEENESLSLEISVSEYVYKIKELDDECFKEKVESE-----L 1277  
 OY 283 TKRISAIKDSG-----ELKAEIERAKKCSSEFTAKL--GEHTDLGKEGYTDDAKKRAIL 335  
 DB 1278 EDKLKASEDKNSKYLDELQKSEKIKELDLAKTTELKIQLEKITLNSAKKESSELSRLK 1337  
 OY 336 KTNNDKTKGAD-LEKI-----FESYKNSLKAKEMLNSVKE 372  
 DB 1338 KTSSEERKNAEOLLEKLNKNEIQIKNOAFEXERKLLNCGSSITTOEYSE 1385

RESULT 11  
 ID GENE\_HUMAN STANDARD; PRT; 3210 AA.  
 AC P49454; Q13246; Q13171;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 20-FEB-1996 (Rel. 33, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CENP-F KINETOCHORE PROTEIN (CENTROMERE PROTEIN F) (MITOSIN) (AH  
 DE ANTIGEN).  
 GN CENPF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Breast carcinoma;  
 RX MEDLINE-95348175; PubMed-7542657;  
 RA Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;  
 RT "CENP-F is a protein of the nuclear matrix that assembles onto  
 RT kinetochores at late G2 and is rapidly degraded after mitosis.";  
 RL J. Cell Biol. 130:507-518(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95379848; PubMed-7651420;  
 RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,  
 RA Jones D., Yang-Feng T.L., Lee W.-H.;  
 RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein  
 RT that is specifically involved in mitotic-phase progression.";  
 RL Mol. Cell. Biol. 15:5017-5029(1995).  
 RN [3]  
 RP SEQUENCE OF 2194-3210 FROM N.A.  
 RX MEDLINE-95336446; PubMed-7612011;  
 RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsker T.A. Jr., Joshi H.C.;  
 RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal  
 RT domain sufficient for nuclear localization.";  
 RL Biochem. Biophys. Res. Commun. 213:220-228(1995).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE-95370296; PubMed-7642639;  
 RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;  
 RT "The C terminus of mitosis is essential for its nuclear localization,  
 RT centromere/kinetochore targeting, and dimerization.";  
 RL J. Biol. Chem. 270:19545-19550(1995).  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE-98437347; PubMed-9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;

RT "Characterization of the kinetochore binding domain of CENP-E reveals  
 RT interactions with the kinetochore proteins CENP-F and hBUBR1.";  
 RT J. Cell Biol. 143:49-63(1998).  
 CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCORE FUNCTION, INVOLVED IN  
 CC CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH  
 CC REINOBLASTOMA PROTEIN (RB), CENP-E AND BUBR1.  
 CC -1- SUBUNIT: HOMO- OR HETERODIMER.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),  
 CC REORGANIZATION TO THE KINETOCORE/CENTROMERE (CORONAL SURFACE OF  
 CC THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.  
 CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.  
 CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U19769; AAA82889.1; -;  
 DR EMBL; U30872; AAA82935.1; -;  
 DR EMBL; U25725; AAA86889.1; -;  
 DR HSSP; P02649; ILE4.  
 DR MIM; 600236; -;  
 KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;  
 KM Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.  
 FT DOMAIN 14 197  
 FT DOMAIN 273 769  
 FT DOMAIN 823 1328  
 FT DOMAIN 1642 1746  
 FT DOMAIN 1862 2987  
 FT DOMAIN 2207 2568  
 FT REPEAT 2207 2386  
 FT REPEAT 2389 2568  
 FT SITE 3015 3032  
 FT CONFLICT 16 16  
 FT CONFLICT 250 250  
 FT CONFLICT 272 272  
 FT CONFLICT 611 611  
 FT CONFLICT 1494 1589  
 FT CONFLICT 1611 1611  
 FT CONFLICT 1811 1811  
 FT CONFLICT 2242 2243  
 FT CONFLICT 2335 2335  
 FT CONFLICT 2492 2492  
 FT CONFLICT 2545 2561  
 SQ SEQUENCE 3210 AA; 367589 MW; 11D83324960EA334 CRC64;  
 Query Match 8.8%; Score 162.5; DB 1; Length 3210;  
 Best Local Similarity 22.0%; Pred. No. 4;  
 Matches 84; Conservative 56; Mismatches 147; Indels 95; Gaps 12;  
 Oy 44 KEVALLSSIDIAKAIG---KRIHONNGDTEVHNHNSLAGVAISTLIKOKDLSK 100  
 Db 2208 EEVHQLRGIEKLRVRIEADKKOHLAKLKERRENDLS-----KKEVLE 2256  
 Oy 101 N-----EGLKEKIDAAKCEFTFNKLEKHTDGLKEGVTDADAEAIL 144  
 Db 2257 RELQSENOELVILDAENSKAEVETLTKQLEMAKSLKVELD---VLRSEKELUT 2312  
 Oy 145 KTNQKRTGAEELKLFESVEVL-----SKAAKEMLANSVELTSPVAES 190  
 Db 2313 KOIQEKQOSELDELKLSFSLKEEKQAEIQIKESKTAVENTMQOLKELNEAV---- 2368  
 Oy 191 PAMVNSCKDQNTSANSDESVAKGNPLEISKTTESNAVLAKEVEITLITSDLELAKA 250  
 Db 2369 -----AALGQDQIKMATQSLDPLIEEHQLRNLSIEKLRARLEADKKQOLC 2415  
 Oy 251 IGGKIKNDVSLDNDADHNGSLISGAVYLSNL-----ITKKISAIKDSGELAKAEIKAK 303

Db 2416 VLQOLK-----ESEHNDLKGK--VENLELEIARTNOEHALLENSKGEVETIK 2466  
 Oy 304 KCEEFKAKLKGHTDL-----GKEGVTDNAK-----AIIKTNNDKYGAELEKL 351  
 Db 2467 AKLEGMTQSLRGLELDVYVIRSEKEDLTNLEKQERISELEINVSFENIIQERQEKV 2526  
 Oy 352 FESVKNLSRAKEMELTNSVKEL 373  
 Db 2527 --QMKKSSTAMEMLOQLKEL 2546  
 RESULT 12  
 MS2\_DICDI  
 ID MS2\_DICDI STANDARD; PRT; 2116 AA.  
 AC P08799;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MYOSIN II HEAVY CHAIN, NON MUSCLE.  
 GN MHCA.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellidae; Dictyostellum.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87092266; PubMed=3540939;  
 RA Warrick H.M., de Lozanne A., Leitwand L.A., Spudich J.A.;  
 RT "Conserved protein domains in a myosin heavy chain gene from  
 RL Dictyostelium discoideum.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).  
 RP [2]  
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.  
 RC STRAIN=AX2;  
 RX MEDLINE=90353583; PubMed=2387408;  
 RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,  
 RT Gerisch G.;  
 RL "Replacement of threonine residues by serine and alanine in a  
 RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";  
 RN FEBS Lett. 269:239-243(1990).  
 RP [3]  
 RP PHOSPHORYLATION SITES.  
 RX MEDLINE=88112226; PubMed=2828113;  
 RA Magle G., Noegel A., Scheel J., Gerisch G.;  
 RT "Phosphorylation of threonine residues on cloned fragments of the  
 RL Dictyostelium myosin heavy chain.";  
 RN FEBS Lett. 227:71-75(1988).  
 RP [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
 RX MEDLINE=95345066; PubMed=7619795;  
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,  
 RT Rayment I.;  
 RL "X-ray structures of the myosin motor domain of Dictyostelium  
 RT biochemically complexed with MgADP, Berx and MgADP.ALf4-";  
 RN Biochemistry 34:8960-8972(1995).  
 RP [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.  
 RX MEDLINE=95345067; PubMed=7619796;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the  
 RL truncated head of Dictyostelium discoideum myosin to 2.7-A  
 RT resolution.";  
 RN Biochemistry 34:8973-8981(1995).  
 RP [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RX MEDLINE=96206189; PubMed=8611530;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the  
 RL Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";  
 RN Biochemistry 35:5404-5417(1996).  
 RP [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
 RX MEDLINE=97452580; PubMed=9305951;

RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
 RT "X-ray structures of the MgADP, MgATPgammAS, and MgAMPNP complexes  
 of the Dictyostelium discoidium myosin motor domain.";  
 RL Biochemistry 36:11619-11628(1997).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RX MEDLINE-98070605; PubMed-9405148;  
 RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
 RT "X-ray crystal structure and solution fluorescence characterization  
 of Mg<sub>2</sub>(3')-O-(N-methylanthraniloyl) nucleotides bound to the  
 Dictyostelium discoidium myosin motor domain.";  
 RL J. Mol. Biol. 274:394-407(1997).  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
 ACTIVITY THAT IS ACTIVATED BY ACTIN.  
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI  
 LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
 CORTEX.  
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEMOROSIN (LMM) AND 1 HEAVY MEMOROSIN (HMM). IT CAN BE FURTHER  
 SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
 THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE  
 ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
 POSITION (688).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M14628; AAA33227.1; -  
 DR PIR: A2655; A2655.  
 DR PIR: S00250; S00250.  
 DR PDB: 1MMA: 03-DEC-97.  
 DR PDB: 1HMD: 17-AUG-96.  
 DR PDB: 1MMG: 03-DEC-97.  
 DR PDB: 1MMN: 03-DEC-97.  
 DR PDB: 1MNE: 17-AUG-96.  
 DR PDB: 1VOM: 23-DEC-96.  
 DR PDB: 1LVK: 28-JAN-98.  
 DR Dictydb: DD01008; mbca.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00242; MISC; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR Myosin: Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation;  
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.  
 FT DOMAIN 762 791 IQ.  
 FT NP\_BIND 817 2116 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP.  
 FT DOMAIN 638 660 ACTIN-BINDING.  
 FT DOMAIN 738 752 ACTIN-BINDING.  
 FT MOD\_RES 130 130 METHYLATION (DI-) (POTENTIAL).

FT MOD\_RES 678 678 ALKYLATION (SH-1).  
 FT MOD\_RES 1823 1823 PHOSPHORYLATION (BY MCK).  
 FT MOD\_RES 1833 1833 PHOSPHORYLATION (BY MCK).  
 FT MOD\_RES 2029 2029 PHOSPHORYLATION (BY MCK).  
 SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;  
 Query Match 8.6%; Score 159; DB 1; Length 2116;  
 Best Local Similarity 21.2%; Pred. No. 3.6;  
 Matches 98; Conservative 73; Mismatches 153; Indels 138; Gaps 18;  
 QY 9 DGMTSANSADSVK-----PMLTEISKITPTSNVLAKEVEALLSIDELAIAKIG 62  
 DB 964 DGSDITISREKIKDELQKEVEELTESFSSEKDKGL--EKTRVRLQSELDIVR--- 1018  
 QY 63 KKHQNNGLDTEYNHNSLAGAVALSTLIKRLDKGNGLKEIKDAKCKSEFTTNKL 122  
 DB 1019 -----LDSETKDKSELLRQKKLEELKQVQALAE-----TAAKLAQEAANKKL 1064  
 QY 123 KEKHTDLGKEGVTADAKKAILKTNGT-----KTGA----- 154  
 DB 1065 QGEYTELNEKFNSEVARSNVEKSKKLTLESQVAVNNNEDEEKKNRDALKKKKALDAML 1124  
 QY 155 EELGKLFESE-----VEVLSKAKEMLAIVKVELTSPVY----- 187  
 DB 1125 EEKKDLESTGEGKKSLYLDKVAQESDMELRNOISELOSTIKKEIKISTLEGVARLQ 1184  
 QY 188 ---AESPAMVNNSGKD-----GNTSANSADSVKGPMLTEISKITPSNAVLAKE 236  
 DB 1185 GELAAQGLAKSNVEKQKKVELDLEKSNQALAEETAKQAKDLKKRKE-----LSE 1238  
 QY 237 VETLLTSLDELAKAIGKKI-----NDVSLNEADNGSLISGAVLISNLT 284  
 DB 1239 VQY-----QLSEANNKNVNSDSTNKHLETSPFNKLLEAEQAK-----QALEKK 1284  
 QY 285 KI---SAIKDSGLKAEIKAKKCEPFAKLGHEHTDGLKEGVTDDA--KKAILTND 340  
 DB 1285 RLGESEELKHNVDLEEKQKSNKRYDLEKVEL--KQDIEEVASKKAATVTEKK 1343  
 QY 341 KTKGADELEKLF-----ESVKNLS--KAAKEMLTNSVKE 372  
 DB 1344 KESELDIEIKRYADVVSRRKSVQELTKLQAKKEELRNTAE 1365  
 RESULT 13  
 LMB1\_MOUSE STANDARD; PRT; 1786 AA.  
 AC P02469;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).  
 GN LAMB-1 OR LAMB-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-87147212; PubMed-3493487;  
 RA Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;  
 RT "Sequence of the cDNA encoding the laminin B1 chain reveals a  
 multidomain protein containing cysteine-rich repeats.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).  
 RN [2]  
 RP SEQUENCE OF 1292-1786 FROM N.A.  
 RX MEDLINE-85051302; PubMed-6209134;  
 RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;  
 RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of  
 coiled-coil alpha-helix.";  
 RL EMBO J. 3:2355-2362(1984).  
 RN [3]  
 RP SEQUENCE OF 165-172; 539-547 AND 712-719.

FT	DOMAIN	1398	1430	DOMAIN ALPHA.
FT	DOMAIN	1431	1786	DOMAIN I.
FT	DOMAIN	1216	1315	COILED COIL (POTENTIAL).
FT	DOMAIN	1368	1378	COILED COIL (POTENTIAL).
FT	DOMAIN	1448	1778	COILED COIL (POTENTIAL).
FT	DISULFID	271	280	BY SIMILARITY.
FT	DISULFID	273	298	BY SIMILARITY.
FT	DISULFID	300	309	BY SIMILARITY.
FT	DISULFID	312	332	BY SIMILARITY.
FT	DISULFID	335	344	BY SIMILARITY.
FT	DISULFID	337	362	BY SIMILARITY.
FT	DISULFID	365	374	BY SIMILARITY.
FT	DISULFID	377	395	BY SIMILARITY.
FT	DISULFID	398	411	BY SIMILARITY.
FT	DISULFID	400	426	BY SIMILARITY.
FT	DISULFID	428	437	BY SIMILARITY.
FT	DISULFID	440	455	BY SIMILARITY.
FT	DISULFID	458	472	BY SIMILARITY.
FT	DISULFID	460	479	BY SIMILARITY.
FT	DISULFID	481	490	BY SIMILARITY.
FT	DISULFID	493	507	BY SIMILARITY.
FT	DISULFID	773	785	BY SIMILARITY.
FT	DISULFID	775	792	BY SIMILARITY.
FT	DISULFID	794	803	BY SIMILARITY.
FT	DISULFID	806	818	BY SIMILARITY.
FT	DISULFID	821	833	BY SIMILARITY.
FT	DISULFID	823	840	BY SIMILARITY.
FT	DISULFID	842	851	BY SIMILARITY.
FT	DISULFID	854	864	BY SIMILARITY.
FT	DISULFID	867	876	BY SIMILARITY.
FT	DISULFID	869	883	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	898	914	BY SIMILARITY.
FT	DISULFID	917	933	BY SIMILARITY.
FT	DISULFID	919	944	BY SIMILARITY.
FT	DISULFID	946	955	BY SIMILARITY.
FT	DISULFID	958	973	BY SIMILARITY.
FT	DISULFID	976	990	BY SIMILARITY.
FT	DISULFID	978	997	BY SIMILARITY.
FT	DISULFID	1000	1009	BY SIMILARITY.
FT	DISULFID	1012	1025	BY SIMILARITY.
FT	DISULFID	1084	1096	BY SIMILARITY.
FT	DISULFID	1086	1103	BY SIMILARITY.
FT	DISULFID	1105	1114	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1132	1144	BY SIMILARITY.
FT	DISULFID	1134	1151	BY SIMILARITY.
FT	DISULFID	1153	1162	BY SIMILARITY.
FT	DISULFID	1165	1176	BY SIMILARITY.
FT	DISULFID	1179	1187	INTERCHAIN (PROBABLE).
FT	DISULFID	1182	1192	INTERCHAIN (PROBABLE).
FT	DISULFID	1185	1785	INTERCHAIN (PROBABLE).
FT	CARBOHYD	120	120	N-LINKED (GLCNAC. .)
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. .)
FT	CARBOHYD	519	519	N-LINKED (GLCNAC. .)
FT	CARBOHYD	677	677	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1041	1041	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1195	1195	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1279	1279	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1336	1336	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1343	1343	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1487	1487	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1533	1533	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1542	1542	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1643	1643	N-LINKED (GLCNAC. .)
FT	CONFLICT	1531	1534	SGNA -> MEMB (IN REF. 2).
FT	CONFLICT	1749	1749	D -> N (IN REF. 2).
50	SEQUENCE	1786	196504	MM; 846671B7BF41AA74 CRC64;



RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.  
 RA HEMLINE-96419133; PubMed=805510;  
 RT "Structure of the regulatory domain of scallop myosin at 2-A  
 resolution: implications for regulation."  
 RL Structure 4:21-32(1996).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X55714; CAA39247.1;  
 DR PIR; S13557; S13557.  
 DR PIR; A40997; A40997.  
 DR PDB; 1SCM; 30-APR-94.  
 DR PDB; 1WDC; 11-JUL-96.  
 DR InterPro; IPR000048; IQ.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR InterPro; IPR002017; Spectrin.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00576; myosin\_head; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.  
 FT DOMAIN 1 777  
 FT DOMAIN 778 805  
 FT DOMAIN 806 835  
 FT DOMAIN 836 1938  
 FT NE\_BIND 176 183  
 FT MOD\_RES 693 693  
 FT MOD\_RES 703 703  
 FT MOD\_RES 778 821  
 FT HELIX 822 823  
 FT TURN 825 833  
 FT TURN 834 835  
 SQ SEQUENCE 1938 AA; 222821 MW; A5CCE4127D1A4896 CRC64;  
 Query Match 8.5%; Score 157; DB 1; Length 1938;  
 Best Local Similarity 22.28; Pred. No. 4;  
 Matches 87; Conservative 61; Mismatches 166; Indels 78; Gaps 13;  
 OY 18 DESVAGPMLTEISKRTSDNSAVLAVKEVEALLSSIDEIAAKAGKIRHONGIDT-----73  
 DB 925 DEEDAAADLEGIKKMEADNA-----NLKDDIDLENTLOKAFOAKAHKNDQISTLOGE 978  
 OY 74 ---EYNHNGSLAGAYASTLIKOKLDLKNGLKERIDAKKCSFTTNKLEKHHDLG 130  
 DB 979 ISQDEHGHGKLNKERRKALEANKKTSLSQAE--EDKCNHLNKKAKALEQALDELDELNLE 1036  
 OY 131 KEGYTDADAKAEALKTNGTGTGKAEEELGKLFEESEYVLSKAKEMLANSVKELTSPYVAES 190  
 DB 1037 REKKVRDVERA-----KRKVEQDLKSTQENVEDLEVRKRE-LEENVRRKEAEI-----1084

OY 191 PAMVNGSGKQONTANSADSVK--GPNLTETISKRT--TESNAVVLAVKEVEETLTSIDEL 247  
 DB 1085 -SINSLEDEQNLVSOLORRIRKELQARIELEEELEAEERNARAKVEKQRAELNELEEL 1143  
 OY 248 AKAIK-----KIKNDV---SLDNEADHNGSLISGATVLSNLTGK 284  
 DB 1144 GERLDEAGATSAQIELNKKREAEELKIRPDLFEASLQHEAO-----ISALRRK 1192  
 OY 285 KISAKDSGLKAEIEKAKKCSFTTAKLKGHTDGLKGEVTDNNAKKAAILKTNNDXTKG 344  
 DB 1193 HODANEMADQVDLOLQVKRSKLEKDKDLKREMDLESQ-----MTHMKMKNG 1240  
 OY 345 ADE-LKLFESYVNLKSKAKEMLTNSVKELTS 375  
 DB 1241 CSEKVKQFESQMSDLNARLEDSQRSINELQS 1272

Search completed: March 18, 2002, 10:11:42  
 Job time: 960 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds  
(without alignments)  
153.266 Million cell updates/sec

Title: US-09-596-746A-28

Perfect score: 1859  
Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212257 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/6C\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	968.5	52.1	466	4	US-08-235-836C-110 Sequence 110, App
2	945	50.8	210	1	US-08-158-353-3 Sequence 3, Appl
3	938	50.5	210	4	US-08-209-603E-15 Sequence 15, Appl
4	933	50.2	209	4	US-09-196-293-15 Sequence 15, Appl
5	932	50.1	210	4	US-08-235-836C-30 Sequence 30, Appl
6	929	50.0	210	4	US-08-235-836C-122 Sequence 122, App
7	915	49.2	466	4	US-08-235-836C-107 Sequence 107, App
8	685	36.8	212	1	US-08-158-353-2 Sequence 2, Appl
9	670	36.0	212	1	US-08-031-295-2 Sequence 2, Appl
10	670	36.0	212	4	US-07-903-580-2 Sequence 2, Appl
11	631	33.9	212	1	US-08-158-353-4 Sequence 4, Appl
12	629	33.8	212	4	US-09-196-293-11 Sequence 11, Appl
13	629	33.8	212	4	US-08-209-603E-11 Sequence 11, Appl
14	626	33.7	212	4	US-08-235-836C-34 Sequence 34, Appl
15	614.5	33.1	209	4	US-08-235-836C-32 Sequence 32, Appl
16	592.5	31.9	207	4	US-08-235-836C-36 Sequence 36, Appl
17	171	3.2	194	4	US-09-364-083-2 Sequence 2, Appl
18	164	8.8	2482	1	US-08-328-254-6 Sequence 6, Appl
19	163	8.8	3248	1	US-08-353-700-1 Sequence 1, Appl
20	163	8.8	3248	5	PCT-US95-16216-1 Sequence 1, Appl
21	155	8.3	1388	2	US-08-685-576-1 Sequence 1, Appl
22	151.5	8.1	1388	2	US-08-685-576-4 Sequence 4, Appl
23	148	8.0	3111	2	US-08-460-309-4 Sequence 4, Appl
24	148	8.0	3111	2	US-08-125-077-4 Sequence 4, Appl
25	146	7.9	688	3	US-09-141-047-8 Sequence 8, Appl
26	145.5	7.8	1196	1	US-08-144-121-4 Sequence 4, Appl
27	145.5	7.8	1196	2	US-08-735-893-4 Sequence 4, Appl

28	144.5	7.8	1164	4	US-08-923-992A-2 Sequence 2, Appl
29	144	7.7	1713	3	US-08-600-982-24 Sequence 24, Appl
30	144	7.7	1713	5	PCT-US94-10261A-24 Sequence 24, Appl
31	143	7.7	630	4	US-08-973-462-9 Sequence 9, Appl
32	142.5	7.7	1164	4	US-08-923-992A-10 Sequence 10, Appl
33	142	7.6	1507	3	US-08-929-329-5 Sequence 5, Appl
34	142	7.6	1786	4	US-08-973-462-8 Sequence 8, Appl
35	139.5	7.5	1588	5	PCT-US93-07261-11 Sequence 11, Appl
36	139.5	7.5	1653	5	PCT-US93-07261-16 Sequence 16, Appl
37	138	7.4	1098	4	US-08-923-992A-8 Sequence 8, Appl
38	137	7.4	641	4	US-08-961-083-160 Sequence 160, App
39	136.5	7.3	1038	4	US-09-541-782-4 Sequence 4, Appl
40	135.5	7.3	1104	4	US-08-923-992A-4 Sequence 4, Appl
41	135.5	7.3	1128	4	US-08-923-992A-6 Sequence 6, Appl
42	135	7.3	573	4	US-08-235-836C-112 Sequence 112, App
43	134	7.2	376	6	Patent No. 5180810-1 Sequence 23, Appl
44	132	7.1	1561	3	US-08-894-017-23 Sequence 23, Appl
45	131.5	7.1	695	1	US-08-127-499A-23 Sequence 23, Appl

## ALIGNMENTS

RESULT 1  
US-08-235-836C-110  
Sequence 110, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110  
Query Match 52.1% Score 968.5; DB 4; Length 466;  
Best Local Similarity 59.5% Pred. No. 5.2e-62;  
Matches 229; Conservative 32; Mismatches 93; Indels 31; Gaps 9;  
OY 1 MACNNSGKDGTSANSADESVKGNLPEINKKITDSNVLAVKEVALLSIDIAKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDETIAKA 76  
QY 61 ICKRIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEDIKDAKCKSEFTN 120  
Db 77 ICKRIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEDIKDAKCKSEFTN 120  
QY 121 KIKRKHDTLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
Db 137 KIKRKHDTLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
QY 181 ELTSPVVAESPCKPKRQ-VNSSLID--EKNSVYDLPGEMKVLKSKKKNODGYTDIATVD 252  
Db 197 ELTSPVVAESPCKPKRQ-VNSSLID--EKNSVYDLPGEMKVLKSKKKNODGYTDIATVD 252  
QY 239 EVETLTS-----IDLAKAIKRIKNDVSLDNEADHNSISAVLI--SNLITKRI 289  
Db 253 KIKRKHDTLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
QY 290 SAIKDSGELKAEIKRACCKSEFTAKLGEHTDLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
Db 308 TS-KDKSSTEKEFNEKEVESEKIIITRADG--TRLEVTGKISDGSKAKKAVLAKGYVLEGL 364  
QY 350 ELEKLFESV-----KMLSKAK 366  
Db 365 TAEKTLVVEGTVLTKNISKSGE 389

RESULT 2  
US-08-158-353-3  
Sequence 3, Application US/08158353  
Patent No. 3620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: DCT93-05  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9340  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-3

Query Match 50.8%; Score 945; DB 1; Length 210;  
Best Local Similarity 97.9%; Pred. No. 8.9e-61;  
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDETIAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDETIAKA 76  
QY 61 ICKRIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEDIKDAKCKSEFTN 120  
Db 77 ICKRIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEDIKDAKCKSEFTN 120  
QY 121 KIKRKHDTLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
Db 137 KIKRKHDTLGEKGYTDADAKKAIILKANGTKRGAEELGKLFESVYLSKAKEMLANSVK 180  
QY 181 ELTSPVVAESPCKPKR 194  
Db 197 ELTSPVVAESPCKPKR 210

RESULT 3  
US-08-209-603E-15  
Sequence 15, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILKE, BETTINA  
APPLICANT: PREAC-MURISIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROOKS HAIDT HAEFFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
CLASSIFICATION: 436  
PRIOR APPLICATION: 436  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 697-5635  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
DESCRIPTION: N/A  
HYPOTHEICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE: B. BURGDOFFER  
ORGANISM: B. BURGDOFFER  
IMMEDIATE SOURCE: DSM 5662  
LIBRARY: DSM 5662

POSITION IN GENOME: N/A  
FEATURE: N/A  
IDENTIFICATION METHOD: amino acid analysis  
PUBLICATION INFORMATION: N/A  
US-08-209-603E-15

Query Match 50.5%; Score 938; DB 4; Length 210;  
Best Local Similarity 97.4%; Pred. No. 2.8e-60;  
Matches 189; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76  
OY 61 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136  
OY 121 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 180  
DB 137 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 196  
OY 181 ELTSPVVAESPKP 194  
DB 197 ELTSPVVAESPKP 210

RESULT 4  
US-09-196-293-15  
Sequence 15, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Soutcheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Praec-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: 738.001US2  
CURRENT APPLICATION NUMBER: US/09/196,293  
CURRENT FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FASTED for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 209  
TYPE: PRP  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-15

Query Match 50.2%; Score 933; DB 4; Length 209;  
Best Local Similarity 97.9%; Pred. No. 6.4e-60;  
Matches 188; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76  
OY 61 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136

OY 121 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 180  
DB 137 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 196  
OY 181 ELTSPVVAESPK 192  
DB 197 ELTSPVVAESPK 208

RESULT 5  
US-08-235-836C-30  
Sequence 30, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-Apr-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-30

Query Match 50.1%; Score 932; DB 4; Length 210;  
Best Local Similarity 96.9%; Pred. No. 7.6e-60;  
Matches 188; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76  
OY 61 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136  
OY 121 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 180  
DB 137 KKEKHTDGLKEGYTDADAKAELIKANGTKTGAELGLKFESEVYLSKAKEMLANSVK 196  
OY 181 ELTSPVVAESPKP 194

Db 197 ELTSPVAESPKRP 210

## RESULT 6

US-08-235-836C-122

Sequence 122, Application US/08235836C

Patent No. 6248562

## GENERAL INFORMATION:

APPLICANT: Dunn, John J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESS: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/235,836C

FILING DATE: 29-APR-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-7329

INFORMATION FOR SEQ ID NO: 122:

SEQUENCE CHARACTERISTICS:

LENGTH: 588 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-122

Query Match

Best Local Similarity 50.0%; Score 929; DB 4; Length 588;

Matches 191; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Db 4 NNSGKGNTSANSADSEYKGNLTKITDSNAVLAVKEVEALLSIDEIAKAIGK 63

Db 286 NNSGKGNTSANSADSEYKGNLTKITDSNAVLAVKEVEALLSIDEIAKAIGK 345

Db 64 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGLKEKIDAAKCEFTTNK 123

Db 346 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGLKEKIDAAKCEFTTNK 405

Db 124 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 183

Db 406 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 465

Db 184 SPVVAESPKRPSMNNSEKDGNTSANSADSEYK 216

Db 466 SPVVAESPKRPGTMAQYQNMHLSNKSASQVNR 498

RESULT 7

US-08-235-836C-107

Sequence 107, Application US/08235836C

Patent No. 6248562

US-08-235-836C-107

Sequence 107, Application US/08235836C

Patent No. 6248562

## GENERAL INFORMATION:

APPLICANT: Dunn, John J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESS: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/235,836C

FILING DATE: 29-APR-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-7329

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-107

Query Match

Best Local Similarity 49.2%; Score 915; DB 4; Length 466;

Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 4 NNSGKGNTSANSADSEYKGNLTKITDSNAVLAVKEVEALLSIDEIAKAIGK 63

Db 276 NNSGKGNTSANSADSEYKGNLTKITDSNAVLAVKEVEALLSIDEIAKAIGK 335

Db 64 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGLKEKIDAAKCEFTTNK 123

Db 346 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGLKEKIDAAKCEFTTNK 395

Db 124 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 183

Db 396 AKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 455

Db 184 SPVVAESPKRP 194

Db 456 SPVVAESPKRP 466

RESULT 8

US-08-158-353-2

Sequence 2, Application US/08158353

Patent No. 5620862

GENERAL INFORMATION:

APPLICANT: Padula, Steven J.

TITLE OF INVENTION: Methods for Diagnosing Early Lyme

TITLE OF INVENTION: Disease

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-2

Query Match 36.8%; Score 685; DB 1; Length 212;  
Best Local Similarity 74.1%; Pred. No. 3.8e-42;  
Matches 146; Conservative 19; Mismatches 28; Indels 4; Gaps 3;

OY 1 MACNNSGKDGNTSANSAPESVYKPNLITEINKKITDSNAVLAVKEVEALLSSIDEIAAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVAGPNLITELSKITTESNAVLAVKEVEALLSSIDEIAAKA 76  
OY 61 IGKKIHQNNGLDTEENNHNGLAGAVYSTLIKOKLDGLKN-EGLEKIDAAKCSSEFT 119  
DB 77 IGNLIAQ-NGLNAGANONGSLAGAVYSTLIAEKLIDGLKNEBELKEKIEDAKKCKAKFT 135  
OY 120 NKLAKKHDTLQ--KEGVTDADAKAAILKANGTKTGAEELGLFESVEVLSAAKEMLAN 177  
DB 136 DKLKSHAEIGANGAATDANAKAAILKTNGTKDGAQELERLFESVKNLSKAAOETLNN 195  
OY 178 SVKELTSPVVAESPCKP 194  
DB 196 SVKELTSPVVAENPKP 212

RESULT 9  
US-08-031-295-2  
Sequence 2, Application US/08031295  
Patent No. 5530103  
GENERAL INFORMATION:  
APPLICANT: LIVER, Ian  
APPLICANT: DORNER, Freidrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/031,295  
FILING DATE: 19930312  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/903,580  
FILING DATE: 25-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/142 IMNU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-031-295-2

Query Match 36.0%; Score 670; DB 1; Length 212;  
Best Local Similarity 73.0%; Pred. No. 4.5e-41;  
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

OY 1 MACNNSGKDGNT-SANSAPESVYKPNLITEINKKITDSNAVLAVKEVEALLSSIDEIAK 59  
DB 17 ISCNNSGKGDSDSSTNPADSAKGNLITELSKITTESNAVLAVKEVEALLSSIDEIAK 76  
OY 60 AIGKKIHQNNGLDTEENNHNGLAGAVYSTLIKOKLDGLKNG-LKEKIDAAKCSSEFT 118  
DB 77 AIGKKIHQNNGLDTEENNHNGLAGAVYSTLIKOKLDGLKNGELKAKIEDAKKCSSEFT 136  
OY 119 TNKLEKHTDLQEGVTDADAKAAILKANGTKTGAEELGLFESVEVLSAAKEMLAN 178  
DB 137 TKLIAGHQLDIGDITDSDKEALIKTNGTKTGAEELVLSAAKEMLAN 196  
OY 179 SVKELTSPVVAESPCKP 194  
DB 197 VKELTSPVVAETPKP 212

RESULT 10  
US-07-903-580-2  
Sequence 2, Application US/07903580  
Patent No. 6221363  
GENERAL INFORMATION:  
APPLICANT: LIVER, Ian  
APPLICANT: DORNER, Freidrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/903,580  
FILING DATE: 19920625  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/131 IMM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-903-580-2

Query Match 36.0%; Score 670; DB 4; Length 212;  
Best Local Similarity 73.0%; Pred. No. 4,56-41;  
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;  
QY 1 MACNNGKRGNT-SANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETIAK 59  
DB 17 ISCNNGSGKGDASTNPADESAAGPNTLEISKTTDSNAVLAVKEVEALLSIDETIAK 59  
QY 60 AIGKKIHQNNGLDTEENNHNHNSLGAVALSTLTKOKIDGLKN-EGLEKIDAKKCESEF 118  
DB 77 AIGKKIQONNGLGANADKNSLGAVALSTLTKOKIDGLKN-EGLEKIDAKKCESEF 118  
QY 119 TNKLEKHTDGRGCVTDADAKKAIKANGTKRGAEEIGKLFESVEVLSKAKEMLANS 178  
DB 137 TKKLAAHQIDGADNDNSKEAIKTKTNGKGAEEIVKLSSEVASLSKAQOASANS 196  
QY 179 VKELTSPVAESPKRP 194  
DB 197 VKELTSPVAETPKRP 212

RESULT 11  
US-08-158-353-4  
Sequence 4, Application US/08158353  
Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-9540  
TELEFAX: 617-861-6240  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-4

Query Match 33.8%; Score 629; DB 1; Length 212;  
Best Local Similarity 69.9%; Pred. No. 2,8e-38;  
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;  
QY 1 MACNNGKRGNT-SANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETIAK 59  
DB 17 ISCNNGSGKGDASTNPADESAAGPNTLEISKTTDSNAVLAVKEVEALLSIDETIAK 59  
QY 60 AIGKKIHQNNGLDTEENNHNHNSLGAVALSTLTKOKIDGLKN-EGLEKIDAKKCESEF 118  
DB 77 AIGKKIQONNGLGANADKNSLGAVALSTLTKOKIDGLKN-EGLEKIDAKKCESEF 118  
QY 119 TNKLEKHTDGRGCVTDADAKKAIKANGTKRGAEEIGKLFESVEVLSKAKEMLANS 178  
DB 137 TKKLAAHQIDGADNDNSKEAIKTKTNGKGAEEIVKLSSEVASLSKAQOASANS 196  
QY 179 VKELTSPVAESPKRP 194  
DB 197 VKELTSPVAETPKRP 212

RESULT 12  
US-09-196-293-11  
Sequence 11, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Sautscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: 738,001US2  
CURRENT APPLICATION NUMBER: US/09/196,293  
EARLIER FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE B39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 33.8%; Score 629; DB 4; Length 212;

```

1 CLASSIFICATION: 435
2
3 PRIOR APPLICATION DATA:
4
5 APPLICATION NUMBER: US 08/148,191
6
7 FILING DATE: 01-11-93
8
9 ATTORNEY/AGENT INFORMATION:
10
11 NAME: Bogosian, Margaret C.
12 REGISTRATION NUMBER: 25,324
13 REFERENCE/DOCKET NUMBER: BUL93-284
14
15 TELECOMMUNICATION INFORMATION:
16
17 TELEPHONE: (516) 282-7338
18
19 TELEFAX: (516) 282-7729
20
21 INFORMATION FOR SEQ ID NO: 34:
22
23 SEQUENCE CHARACTERISTICS:
24
25 LENGTH: 212 amino acids
26
27 TYPE: amino acid

```

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-34

Query Match 33.7%; Score 626; DB 4; Length 212;  
Best Local Similarity 69.4%; Pred. No. 6,3e-38;  
Matches 136; Conservative 16; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSDESVKGNLTLEINKKTTDSNAVLAVKEVEALLSIDETIAK 59  
DB 17 ISCSNSG--GDTASTNPDESAGKPNLTISKRTTDSNAVLAVKEVEALLSIDETIAK 76  
QY 60 AIGKTIHONNGLDTENNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFTF 118  
DB 77 AIGKTIHONNGLDTENNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFTF 136  
119 TNKLEKHTDLGKGVTDADAKKAIILKANGTKTKGAEELGKFESVEVLSKAKEMLANS 178  
DB 137 TNKLSGSHADLGKQDATTDDHAKAAILKTHATTDKGAKEFKDFESVEGLLKAQVALTNS 196  
QY 179 VKELTSPVAESPCKP 194  
DB 197 VKELTSPVAESPCKP 212

## RESULT 15

US-08-235-836C-32  
Sequence 32, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polyptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 209 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-32

Query Match 33.1%; Score 614.5; DB 4; Length 209;  
Best Local Similarity 68.2%; Pred. No. 4.1e-37;

Matches 133; Conservative 20; Mismatches 39; Indels 3; Gaps 2;  
QY 1 MACNSGKDGNTSANSADSVKGNLTLEINKKTTDSNAVLAVKEVEALLSIDETIAK 60  
DB 17 ISCSNSG--GDTASTNPDESAGKPNLTISKRTTDSNAVLAVKEVEALLSIDETIAK 74  
QY 61 IGGKTIHONNGLDTENNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFTF 119  
DB 75 IGGKTIHONNGLDTENNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFTF 134  
QY 120 NKLKHTDLGKGVTDADAKKAIILKANGTKTKGAEELGKFESVEVLSKAKEMLANSV 179  
DB 135 NKLKSHADLGVAATAATDDHAKAAILKSNPTKDKGAKALKDLSSEVSESLAKAAQVALANSV 194  
QY 180 KELTSPVAESPCKP 194  
DB 195 KELTSPVAESPCKP 209

Search completed: March 18, 2002, 09:55:32  
Job time: 345 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:42 : Search time 621.2 Seconds  
(without alignments)  
168.954 Million cell updates/sec

Title: US-09-596-746A-28

Perfect score: 1859  
Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSYKELTS 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

## Database :

Pending Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1859	100.0	378	19	US-09-596-746A-28
2	1859	100.0	378	19	US-09-596-746A-28
3	1834	98.7	401	19	US-09-596-746A-80
4	1834	98.7	402	19	US-09-596-746A-80
5	1810.5	97.4	375	19	US-09-596-746A-24
6	1805.5	97.1	374	19	US-09-596-746A-24
7	1802.5	97.0	398	19	US-09-596-746A-60
8	1802.5	97.0	399	19	US-09-596-746A-60
9	1614.5	86.8	377	19	US-09-596-746A-30

10	1614.5	86.8	377	19	US-09-596-746A-30	Sequence 30, Appl
11	1589.5	85.5	400	19	US-09-596-746A-82	Sequence 82, Appl
12	1589.5	85.5	401	19	US-09-596-746A-82	Sequence 82, Appl
13	1566	84.2	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1561	84.0	373	19	US-09-596-746A-26	Sequence 26, Appl
15	1558	83.8	397	19	US-09-596-746A-62	Sequence 62, Appl
16	1558	83.8	398	19	US-09-596-746A-62	Sequence 62, Appl
17	1533.5	82.5	371	19	US-09-596-746A-34	Sequence 34, Appl
18	1528.5	82.2	370	19	US-09-596-746A-34	Sequence 34, Appl
19	1525.5	82.1	354	19	US-09-596-746A-64	Sequence 64, Appl
20	1525.5	82.1	355	19	US-09-596-746A-64	Sequence 64, Appl
21	1501.5	80.8	385	19	US-09-596-746A-56	Sequence 56, Appl
22	1496.5	80.5	384	19	US-09-596-746A-56	Sequence 56, Appl
23	1494.5	80.4	369	19	US-09-596-746A-40	Sequence 40, Appl
24	1493.5	80.3	408	19	US-09-596-746A-78	Sequence 78, Appl
25	1493.5	80.3	409	19	US-09-596-746A-78	Sequence 78, Appl
26	1489.5	80.1	368	19	US-09-596-746A-40	Sequence 40, Appl
27	1486.5	80.0	392	19	US-09-596-746A-72	Sequence 72, Appl
28	1486.5	80.0	393	19	US-09-596-746A-72	Sequence 72, Appl
29	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
30	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
31	1289	69.3	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284	69.1	369	19	US-09-596-746A-36	Sequence 36, Appl
33	1281	68.9	393	19	US-09-596-746A-66	Sequence 66, Appl
34	1281	68.9	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1250	67.2	368	19	US-09-596-746A-42	Sequence 42, Appl
36	1245	67.0	367	19	US-09-596-746A-42	Sequence 42, Appl
37	1242	66.8	391	19	US-09-596-746A-74	Sequence 74, Appl
38	1242	66.8	392	19	US-09-596-746A-74	Sequence 74, Appl
39	1217	65.5	370	19	US-09-596-746A-32	Sequence 32, Appl
40	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
41	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
42	1212	65.2	369	19	US-09-596-746A-32	Sequence 32, Appl
43	1209	65.0	393	19	US-09-596-746A-68	Sequence 68, Appl
44	1209	65.0	394	19	US-09-596-746A-68	Sequence 68, Appl
45	1205	64.8	401	19	US-09-596-746A-70	Sequence 70, Appl

## ALIGNMENTS

RESULT 1  
US-09-596-746-28  
; Sequence 28, Application US/09596746  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seinoist, Gerald  
; APPLICANT: Dykhuisen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEO ID NO 28  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Ospa Chimera  
US-09-596-746-28

Query Match 100.0%; Score 1859; DB 19; Length 378;  
Best Local Similarity 100.0%; Pred. No. 4, 5e-123;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
QY 1 MACNNSGKDGTSANSADESKGNLTIEINKKITDSNAVLAVKEVALLSIDEIAKA 60

```

Db      1  MACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 60
Oy      61  IGKRIHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 120
Db      61  IGKRIHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 120
Oy      121  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 180
Db      121  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 180
Oy      181  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 240
Db      181  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 240
Oy      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 300
Db      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 300
Oy      301  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 360
Db      301  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 360
Oy      361  LSKAKEMLTNSVKELTS 378
Db      361  LSKAKEMLTNSVKELTS 378

```

## RESULT 2

```

US-09-596-746a-28
Sequence 28, Application US/09596746A
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Selmost, Gerald
APPLICANT: Dykhuzen, Daniel
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: Groups of Borrelia burgdorferi
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FASTSEQ for Windows Version 4.0
LENGTH: 378
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Ospec Chimera
US-09-596-746a-28

```

Query Match  
Best Local Similarity 100.0%; Score 1859; DB 19; Length 378;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1  MACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 60
Db      1  MACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 60
Oy      61  IGKRIHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 120
Db      61  IGKRIHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 120
Oy      121  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 180
Db      121  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 180
Oy      181  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 240
Db      181  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 240

```

```

Db      181  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 240
Oy      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 300
Db      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 300
Oy      301  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 360
Db      301  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 360
Oy      361  LSKAKEMLTNSVKELTS 378
Db      361  LSKAKEMLTNSVKELTS 378

```

## RESULT 3

```

US-09-596-746-80
Sequence 80, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Selmost, Gerald
APPLICANT: Dykhuzen, Daniel
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: Groups of Borrelia burgdorferi
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FASTSEQ for Windows Version 4.0
LENGTH: 401
TYPE: PRF
ORGANISM: Ospec Chimera
US-09-596-746-80

```

Query Match  
Best Local Similarity 98.7%; Score 1834; DB 19; Length 401;  
Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

Oy      2  ACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 61
Db      25  SCNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 61
Oy      62  GKRTHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 121
Db      85  GKRTHONNGLDTEENHNHNSLLAGAYASTILIKOKLDGKNGKLEKIDAKKSCSEFTTN 121
Oy      122  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 181
Db      145  KLEKHTDLDGEGVTADADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSVK 181
Oy      182  ELTSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 241
Db      205  LITSPVAESPCKPSVNNNSGKDGNTSANSADSVKGNLTETISKRITESNAVYLAVKEV 241
Oy      242  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 301
Db      265  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLSNLTITKISAIKDSGELKA 301
Oy      302  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 361
Db      335  EIEKAKCSEFTAKLKGHTDLDGEGVTDDNAKKAILEKTNNDKTKGAELEKLFESVKN 361
Oy      362  SKAKEMLTNSVKELTS 378
Db      385  SKAKEMLTNSVKELTS 401

```

```
RESULT 4
US-09-596-746a-80
; Sequence 80, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 80
; LENGTH: 402
; ORGANISM: ospc Chimera
; US-596-746a-80

Query Match 98.7%; Score 1834; DB 19; Length 402;
Best Local Similarity 98.9%; Pred. No. 2.9e-121;
Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSESVKGNLTETINRKTDSNAVLAVKEVEALLSSIDEIAKAI 61
DB 26 SCNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVEALLSSIDEIAKAI 85
OY 62 GKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCSFTFNK 121
DB 86 GKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCSFTFNK 145
OY 122 LKEKHTDLGEGVTDADAKAELIKANGTKTKAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 146 LKEKHTDLGEGVTDADAKAELIKANGTKTKAEELGKLFESVEVLSKAKEMLANSVKE 205
OY 182 LSPVVAESPKRPSMVNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVE 241
DB 206 LSPVVAESPKRPSMVNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVE 265
OY 242 TLTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTITRKISAIRDSGELKAE 301
DB 266 TLTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTITRKISAIRDSGELKAE 325
OY 302 IEKAKKCSSEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKNL 361
DB 326 IEKAKKCSSEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKNL 385
OY 362 SKAKEMLTNSVKELTIS 378
DB 386 SKAKEMLTNSVKELTIS 402

RESULT 5
US-09-596-746a-24
; Sequence 24, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; OTHER APPLICATION NUMBER: US 60/140,042
```

```
;; PRIOR FILING DATE: 1999-06-18
;; NUMBER OF SEQ ID NOS: 84
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 24
;; LENGTH: 375
;; TYPE: PRP
;; ORGANISM: Artificial Sequence
;; OTHER INFORMATION: Ospc Chimera
US-09-596-746a-24

Query Match 97.4%; Score 1810.5; DB 19; Length 375;
Best Local Similarity 98.1%; Pred. No. 1.2e-119;
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSESVKGNLTETINRKTDSNAVLAVKEVEALLSSIDEIAKA 60
DB 1 MACNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVEALLSSIDEIAKA 60
OY 61 ICKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCSFTFNK 120
DB 61 ICKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAKKCSFTFNK 120
OY 121 KLEKHTDLGEGVTDADAKAELIKANGTKTKAEELGKLFESVEVLSKAKEMLANSVKE 180
DB 121 KLEKHTDLGEGVTDADAKAELIKANGTKTKAEELGKLFESVEVLSKAKEMLANSVKE 180
OY 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVE 240
DB 181 ELTSPVVAES---PAMVNNSGKDGNTSANSADSESVKGNLTETISKTITDSNAVLAVKEVE 237
OY 241 ETLTSLDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTITRKISAIRDSGELKA 300
DB 238 ETLTSLDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTITRKISAIRDSGELKA 297
OY 301 EIEKAKKCSSEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKN 360
DB 298 EIEKAKKCSSEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKN 357
OY 361 LSKAKEMLTNSVKELTIS 378
DB 358 LSKAKEMLTNSVKELTIS 375

RESULT 6
US-09-596-746-24
; Sequence 24, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 24
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
US-09-596-746-24

Query Match 97.1%; Score 1805.5; DB 19; Length 374;
```

Best Local Similarity 98.1%; Pred. No. 2.7e-119;  
Matches 370; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNPLTEINKKTTSDNAVLLAVKEVALLSSIDEIAKAI 61
Db 1 ACNNSGKDGNTSANSADSVKGNPLTEINKKTTSDNAVLLAVKEVALLSSIDEIAKAI 60
QY 62 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 121
Db 61 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 120
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 121 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 180
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 241
Db 181 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 237
QY 242 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 301
Db 238 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 297
QY 302 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 298 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 357
QY 362 SKAKEMLTNSVKELTS 378
Db 358 SKAKEMLTNSVKELTS 374
```

## RESULT 7

```
US-09-596-746-60
; Sequence 60, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746-60
```

Query Match 97.0%; Score 1802.5; DB 19; Length 398;  
Best Local Similarity 97.9%; Pred. No. 4.8e-119;  
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNPLTEINKKTTSDNAVLLAVKEVALLSSIDEIAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGNPLTEISKTTSDNAVLLAVKEVALLSSIDEIAKAI 84
QY 62 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 121
Db 85 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 144
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 145 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 204
```

```
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 241
Db 205 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 261
QY 242 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 301
Db 262 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 321
QY 302 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 322 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 381
QY 362 SKAKEMLTNSVKELTS 378
Db 382 SKAKEMLTNSVKELTS 398
```

## RESULT 8

```
US-09-596-746A-60
; Sequence 60, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746A-60
```

Query Match 97.0%; Score 1802.5; DB 19; Length 399;  
Best Local Similarity 97.9%; Pred. No. 4.8e-119;  
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNPLTEINKKTTSDNAVLLAVKEVALLSSIDEIAKAI 61
Db 26 SCNNSGKDGNTSANSADSVKGNPLTEISKTTSDNAVLLAVKEVALLSSIDEIAKAI 85
QY 62 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 121
Db 86 GKTIHONNGLDTEYNNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAAKCSSEFTFNK 145
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 146 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 205
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 241
Db 206 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNPLTEISKTTESNAVYLAKEVE 262
QY 242 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 301
Db 263 TLTSIDELAKAIGKTKINDVSLDNEADHNSLISGAVILSNLITRKISAIKDSGELKAE 322
QY 302 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 323 IEKAKCSEFTAKLGEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 382
QY 362 SKAKEMLTNSVKELTS 378
Db 383 SKAKEMLTNSVKELTS 399
```

```
RESULT 9
US-09-596-746-30
; Sequence 30, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match
Best Local Similarity 86.8%; Score 1614.5; DB 19; Length 377;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDETIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDETIAKA 60
OY 61 IGGKIHQNNGLDTEENNHNSSLGAYASTLIKQKLDGKLNKGLKEKIDAKKCEFTFN 120
DB 61 IGGKIHQNNGLDTEENNHNSSLGAYASTLIKQKLDGKLNKGLKEKIDAKKCEFTFN 120
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
OY 181 ELTSPVVAESPCKPRSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEV 240
DB 181 ELTSPVVAESPCKPRSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEV 240
OY 241 ETLTSTIDELA-KAIGKRIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
DB 241 ETLTSTIDELA-KAIGKRIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
OY 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNAKKAAILKTNDKTKGADELKLFESVK 359
DB 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNAKKAAILKTNDKTKGADELKLFESVK 359
OY 301 EKEENNAKCSDEFTKLEGGHAGLGIENYTDENAKKAAILTTDAKDKGALEKLFRAVE 360
DB 301 EKEENNAKCSDEFTKLEGGHAGLGIENYTDENAKKAAILTTDAKDKGALEKLFRAVE 360
OY 360 NLSKAAKEMLANSVKEL 376
DB 361 NLSKAAKEMLANSVKEL 377

RESULT 10
US-09-596-746a-30
; Sequence 30, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
```

```
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30

Query Match
Best Local Similarity 86.8%; Score 1614.5; DB 19; Length 377;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDETIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDETIAKA 60
OY 61 IGGKIHQNNGLDTEENNHNSSLGAYASTLIKQKLDGKLNKGLKEKIDAKKCEFTFN 120
DB 61 IGGKIHQNNGLDTEENNHNSSLGAYASTLIKQKLDGKLNKGLKEKIDAKKCEFTFN 120
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
OY 181 ELTSPVVAESPCKPRSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEV 240
DB 181 ELTSPVVAESPCKPRSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAVKEV 240
OY 241 ETLTSTIDELA-KAIGKRIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
DB 241 ETLTSTIDELA-KAIGKRIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
OY 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNAKKAAILKTNDKTKGADELKLFESVK 359
DB 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNAKKAAILKTNDKTKGADELKLFESVK 359
OY 301 EKEENNAKCSDEFTKLEGGHAGLGIENYTDENAKKAAILTTDAKDKGALEKLFRAVE 360
DB 301 EKEENNAKCSDEFTKLEGGHAGLGIENYTDENAKKAAILTTDAKDKGALEKLFRAVE 360
OY 360 NLSKAAKEMLANSVKEL 376
DB 361 NLSKAAKEMLANSVKEL 377

RESULT 11
US-09-596-746-82
; Sequence 82, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 400
; TYPE: PRF
; ORGANISM: OspC Chimera
US-09-596-746-82
```

Query Match 85.5%; Score 1589.5; DB 19; Length 400;  
 Best Local Similarity 86.7%; Pred No. 5.7e-104;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEIKRTTDSNAVLAVKEVEALLSIDEIAAKAI 61  
 DB 25 SCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIAAKAI 84  
 QY 62 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTNK 121  
 DB 85 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTNK 144  
 QY 122 LKEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVKE 181  
 DB 145 LKEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVKE 204  
 QY 182 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 241  
 DB 205 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 264  
 QY 242 TLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKISAIKDSGELKA 300  
 DB 265 TLASIDELATKAIGKKIQONGGLAVAGHNGTLAGATITSLITKLDGLKNGELKE 324  
 QY 301 EIEKAKKSEFTAKLKGHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVKN 360  
 DB 325 KIENAKKSEDFTKKLEGEHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVEN 384  
 QY 361 LSKAAKEMLANSVKEL 376  
 DB 385 LAKAAKEMLANSVKEL 400

RESULT 12  
 US-09-596-746a-82  
 ; Sequence 82, Application US/09596746a  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dattwyler, Raymond J.  
 ; APPLICANT: Selinost, Gerald  
 ; APPLICANT: Dykhuitzen, Daniael  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596/746a  
 ; PRIOR FILING DATE: 2000-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 82  
 ; LENGTH: 401  
 ; TYPE: PRF  
 ; ORGANISM: ospc Chimera  
 US-09-596-746a-82

Query Match 85.5%; Score 1589.5; DB 19; Length 401;  
 Best Local Similarity 86.7%; Pred No. 5.8e-104;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEIKRTTDSNAVLAVKEVEALLSIDEIAAKAI 61  
 DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIAAKAI 85  
 QY 62 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTNK 121  
 DB 86 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTNK 145  
 QY 122 LKEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVKE 181

DB 146 LKEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVKE 205  
 QY 182 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 241  
 DB 206 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 265  
 QY 242 TLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKISAIKDSGELKA 300  
 DB 266 TLASIDELATKAIGKKIQONGGLAVAGHNGTLAGATITSLITKLDGLKNGELKE 325  
 QY 301 EIEKAKKSEFTAKLKGHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVKN 360  
 DB 326 KIENAKKSEDFTKKLEGEHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVEN 385  
 QY 361 LSKAAKEMLANSVKEL 376  
 DB 386 LAKAAKEMLANSVKEL 401

RESULT 13  
 US-09-596-746a-26  
 ; Sequence 26, Application US/09596746a  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dattwyler, Raymond J.  
 ; APPLICANT: Selinost, Gerald  
 ; APPLICANT: Dykhuitzen, Daniael  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596/746a  
 ; PRIOR FILING DATE: 2000-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 26  
 ; LENGTH: 374  
 ; TYPE: PRF  
 ; ORGANISM: Artificial Sequence  
 ; OTHER INFORMATION: Opc Chimera  
 US-09-596-746a-26

Query Match 84.2%; Score 1566; DB 19; Length 374;  
 Best Local Similarity 85.9%; Pred No. 2.4e-102;  
 Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEIKRTTDSNAVLAVKEVEALLSIDEIAAKA 60  
 DB 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIAAKA 60  
 QY 61 IGKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTN 120  
 DB 61 IGKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCSSETPTN 120  
 QY 121 KLEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVK 180  
 DB 121 KLEKHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVEVLSKAAKEMLANSVK 180  
 QY 181 ELTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 240  
 DB 181 ELTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 240  
 QY 241 ETLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKISAIKDSGELKA 300  
 DB 238 ETLASIDELATKAIGKKIQONGGLAVAGHNGTLAGAYASTLIKOKLDGLKNGELKE 297  
 QY 300 AEIEKAKKSEFTAKLKGHTDGLKEGYTDADAKKEALITKANGTKGADELKLFESVKN 359  
 DB 298 EKIENAKKSEDFTKKLEGEHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVE 357



OY 360 NUSKAKEMLTNSVKEL 376  
DB 358 LAKAKEMLTNSVKEL 374

## RESULT 14

US-09-596-746-26  
Sequence 26, Application US/09596746  
GENERAL INFORMATION:  
APPLICANT: Dattwyler, Raymond J.  
APPLICANT: Selmoost, Gerald  
APPLICANT: Dykhuisen, Dantale  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 26  
LENGTH: 373  
TYPE: PRN  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Oaspc Chimera  
US-09-596-746-26

Query Match 84.0%; Score 1561; DB 19; Length 373;  
Best Local Similarity 85.9%; Pred. No. 5,4e-102;

Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVEALLSIDEIAAKAI 61  
DB 1 ACNNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVEALLSIDEIAAKAI 60  
OY 62 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAKKCEFTTNK 121  
DB 61 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAKKCEFTTNK 120  
OY 122 LKEKHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEVLSKAKEMLTNSVKEL 181  
DB 121 LKEKHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEVLSKAKEMLTNSVKEL 180  
OY 182 LTPPVAAESPKRPMVNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVE 241  
DB 181 LTPPVAAESPKRPMVNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVE 240  
OY 242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 300  
DB 238 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 297  
OY 301 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEV 360  
DB 298 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEV 357  
OY 361 LSKAKEMLTNSVKEL 376  
DB 358 LAKAKEMLTNSVKEL 373

## RESULT 15

US-09-596-746-62  
Sequence 62, Application US/09596746  
GENERAL INFORMATION:  
APPLICANT: Dattwyler, Raymond J.  
APPLICANT: Selmoost, Gerald  
APPLICANT: Dykhuisen, Dantale

APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 62  
LENGTH: 397  
TYPE: PRN  
ORGANISM: Oaspc Chimera  
US-09-596-746-62

Query Match 83.8%; Score 1558; DB 19; Length 397;  
Best Local Similarity 85.6%; Pred. No. 9,6e-102;

Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVEALLSIDEIAAKAI 61  
DB 25 ACNNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVEALLSIDEIAAKAI 84  
OY 62 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAKKCEFTTNK 121  
DB 61 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKKEKIDAKKCEFTTNK 144  
OY 122 LKEKHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEVLSKAKEMLTNSVKEL 181  
DB 145 LKEKHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEVLSKAKEMLTNSVKEL 204  
OY 182 LTPPVAAESPKRPMVNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVE 241  
DB 205 LTPPVAAESPKRPMVNSGKDGNTSANSADSEYKGNLTETINKKTTDSNAVLLAVEVE 261  
OY 242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 300  
DB 262 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 321  
OY 301 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEV 360  
DB 322 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTKGAELGKLFESYEV 381  
OY 361 LSKAKEMLTNSVKEL 376  
DB 382 LAKAKEMLTNSVKEL 397

Search completed: March 18, 2002, 10:08:43  
Job time: 971 sec





```

: Sequence 7 Application US/09974992
: GENERAL INFORMATION:
: APPLICANT: Mathiesen, Marianne J.
: APPLICANT: Thiesen, Michael
: APPLICANT: Holm, Arne
: APPLICANT: Ostergaard, Soren
: TITLE OF INVENTION: Novel OSCP-derived peptide fragments
: FILE REFERENCE: 459-666P
: CURRENT APPLICATION NUMBER: US/09/974,992
: CURRENT FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: 09/180,089
: PRIOR FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: PCT/DK97/00203
: PRIOR FILING DATE: 1997-05-02
: NUMBER OF SEQ ID NOS: 40
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 212
: TYPE: PRT
: ORGANISM: Borrelia afzelii
: US-09-974-992-7

```

[illegible]

```

RESULT 3
15-09-974-992-3
SEQUENCE 3 Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel Oscp-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: 09/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii
15-09-974-992-3

```

Query Match 32.1%; Score 596.5; DB 6; Length 207;  
Best Local Similarity 66.0%; Pred. No. 2.5e-33;  
Matches 128; Conservative 25; Mismatches 33; Indels 3; Gaps 2

QY	1	MANNSGDGMTSAASDAESYKGNULTEINKKTTIDTSNAVLLAAKEVEALLSDEIDEIAANA	60
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
Db	17	ISCNSSG--GGSATNPDESAGPNLVYISKRTIDTSNAFLLAAKEVEALLSIDEL-SKA	73
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
QY	61	IGKRTHONGNGLDTEENHNGSLACAVAI\$PTL IKOKLIDGLKNEGLKREKIDAAKCS\$ETFIN	120
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
Db	74	IGKRTKNGGTIDNANRR\$ELIACAYEISKLITOKLSVLN\$EELKREKIEAKC\$SEKFTY	133
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
QY	121	KLKEKHTDLGKEGVYTDAAKEALIKANGTFTKGA\$EELKLFES\$EVLIS\$AAKEMLANSYK	180
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
Db	134	KLKSH\$EELIGIOSVODDNAAKAILKTKGTOKDGAKLEELFKL\$E\$LSK\$AAQALANSYK	193
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
QY	181	ELT\$PYVAESPKP	194
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :	
Db	194	ELTNPVAESPKP	207

```

RESULT      4
US-10-032-585-7646
; Sequence 7646, Application US/10032585
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-699
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7646
LENGTH: 1881
TYPE: PR1
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1881)..(1881)
OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

```

Query Match	Best Local Similarity	9.2%	Score 171.5	DB 7	Length 1881
Matches 100	Conservative 78	Mismatched 171	Indels 127	Gaps 17	
QY	1	MACNNSGKDGNTSANSADESVKPGNLTETLNKKTTDDSNVLAWEAVEALSSIDEITAAKA	60		
DB	1411	LAALQETKSNLSDTSMTELEKTELEIKRVNLTETATSELTFRQDNNOSTTETETETAKA	1470		
QY	61	IGKKIHQNNGLDTEHHNNSLLAGAAVISTLLKOKIDGLKNF-----GLKEKIDAKKC	114		
DB	1471	LTK-----SSDLEKCGNOKSELDDSKS-----VKSELKNFNRYNOETYSLKDEIEKKOKE	1523		
QY	115	SETFTNKLKEKHTDGLKEGVTADADAKKALKA-----NG	148		
DB	1524	IYTLQETELKDRISVEYKERAPMLSENSEYTIKREYDOKIKLSKINSIKENHSKEITTHNE	1583		
QY	149	TYTKGAEEGLKGFEEVEVLSKAAKEMLANSVKELTSPVAAESPCKPSPMVNNSGKDGNSTA	208		
DB	1584	QKTSLKQDIAKLSQHE-SAQTDLDEKQMLK-----KAST-----EKHNTESA	1628		
QY	209	NSADESVKPGNLTETLSKKI-----TESNAVLAWEVEITLLSIDELAKKIC	256		
DB	1629	TSIEE--KNNOIKELSEITIKSLTLEKLSGDAKNSQKEFYTLTKKNSDTSKLEKOLEE	1686		
QY	257	--KINDVSLDNEADHNSLISGAVLISNLITTKISAIKDSGELKA-----ETEKAK--	306		
DB	1687	LEKVSQDLQTADE-KLKITERELAKSELEFYVKNNGSLTSELAALTKVYSLEKEEE	1745		
QY	307	-----KCSSEFTALKGEHTDGLK--EGVD-----DNKAKALKTKNNDKTK	346		
DB	1746	LQPLSGNKSKELEDYCK-----HSDISKALKALTDLEKTKTQPDQSKTKTJTELENDLTS	1801		

OY 347 GADELE-----KLFESYKNSKAKEMLTNSYKELTS 378  
DB 1802 TKKELETEKTOTSKFRNLEERKDKETIVKLNKELELLKNDNSGAKKELSEKVS KLES 1857

## RESULT 5

US-10-072-851-15590  
; Sequence 15590, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Cair, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Yamamoto, John D.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jlang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C  
; TITLE OF INVENTION: Proliferation  
; FILE REFERENCE: ELITRA.028A  
; CURRENT APPLICATION NUMBER: US/10/072.851  
; CURRENT FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: fastseq for windows version 4.0  
; SEQ ID NO 15590  
; LENGTH: 1881  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1881)..(1881)  
; OTHER INFORMATION: X-any amino acid  
US-10-072-851-15590

Query Match 9.2%; Score 171.5; DB 7; Length 1881;  
Best Local Similarity 21.0%; Pred. No. 0.0025;  
Matches 100; Conservative 78; Mismatches 171; Indels 127; Gaps 17;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETNKKITDSNAVLAVKEVEALLSIDEIATAKA 60  
DB 1411 LAAOLETKTSNDSTTMELEKTELEKRVNELTEATSELTGLQDNOSLTETIEIEKTKAA 1470  
OY 61 ICKKIHONNGLDTEENNHNHSLAGAVASTLTKOKLDGKNE-----GLKKEIDAKKAC 114  
DB 1471 LTK-----SSKDLVCGNOKSELDDLSK---VKSELKNENKYNQDETTLKDIETEKKE 1523  
OY 115 SETFTNKLKEKHTDLGKECVTDADAKKALKA-----NG 148  
DB 1524 IYTLQTELDKDRISVEVEKEKAMLSENSEYIVKEYSOKIKLSIKENHSKEITTHNB 1583  
OY 149 TTKGABELGKLPESVEYVLSKAKKEMLANSVKELTSPVVAESPKPSMYNNSGKDGNTSA 208  
DB 1584 OKTSLKODIAKLSODHE--SAOTQLEDKEMQLKEL-----KASL-----EKHNTESA 1628  
OY 209 NDAESVKGPNLTETSKT-----TESNAVVLAVKREVEITLLSIDELAKAIGK----- 256  
DB 1629 TSTIEE--KNNQKELSETIKSLTELKTSGDALKOSOKYKTLKTKYNSPTESKLEKOLE 1686  
OY 257 --KIKNDVSLNDADHNGSLGAVLISMLTKKISAIKDSGELKA-----EIEKAK-- 306  
DB 1687 LKVKVSDLOTADE-KLKGITREIALKSELETIVKNSGLSTJSELMALTKVKSLEKEKEE 1745  
OY 307 -----KCEBFTAKLKGHTDLGK--EGVTD-----DNAKAILKTNNDKTK 346

DB 1746 LQFLSGNKKKELEDYIYK-----HSDISEKLAALTDLKEKTKQFDOSKKRKLFLDNDLTS 1801  
OY 347 GADELE-----KLFESYKNSKAKEMLTNSYKELTS 378  
DB 1802 TKKELETEKTOTSKFRNLEERKDKETIVKLNKELELLKNDNSGAKKELSEKVS KLES 1857

## RESULT 6

US-09-708-427-15044  
; Sequence 15044, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708.427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15044  
; LENGTH: 1313  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1313  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1313  
; OTHER INFORMATION: Ceres Seq. ID 1828627  
US-09-708-427-15044

Query Match 9.2%; Score 171; DB 6; Length 1313;  
Best Local Similarity 21.8%; Pred. No. 0.0016;  
Matches 115; Conservative 61; Mismatches 148; Indels 204; Gaps 21;

OY 8 KDGNTSANSADSVKGNLTETNKKITDSNAV-----LLAVKEVEALLSIDEIATAKA 60  
DB 6 KTGMEPTPSKSPPPRSLSKLSASKSDNSASPKVHSRLVKGTETQLOLQIOEDLKKA 65  
OY 61 ICKKIHONNGLDTEENNHNHSLAGAVASTLTKOKL-----LDGLK-----NEGKKEKI 108  
DB 66 -----DEQ-----IELKDKAKAIDDLSESKLVEANEKLEAL 101  
OY 109 DAACKSETF-----TNK-----LKEKHT-DLG----- 130  
DB 102 AAKKRAESEEVEKFRVALEQAGLEAVOKKDYTSKNELESINSQHALDISALLSTTEEL 161  
OY 131 ---KGVTDADAKKAILKANGTKTGAE-----ELGKL-----FES 163  
DB 162 QRYKHLSMTADAKKNAALSHAEATGKIAELHAKEILASSELGRALKALGSKKEKAEIG 221  
OY 164 VEYLSKAKE-----MLANSVKE-----LTSPPVAESP 191  
DB 222 NEVVSXKLEIELLRLGELKSVSLSESLKEQELVOLKVDLEPAKMAESCTNSVVEWK 281  
OY 192 KPSMYNNSGKDGNTSANSADSVKGNLTETNKKITDSNAV-----LA 236  
DB 282 NKVHELEKEVEESNRSKSSSESME-----SVKQALAEIHNHVLHETKSDNAAKKEKIEL 336  
OY 237 VKEVEITLTSIDE-----LAKAIGKIRNDV-----SLDNEADHNGSL 274  
DB 337 EKTIEAQRFDELYEGQVCAKEAEASKLELVLSIKSELEISOEKTRALDNNKKAFTSN- 395  
OY 275 ISGAVLISMLTKKISAIKDSGELKAEIEKAKKCEFTAKLKGHTDLGKEVTDNNAK 334  
DB 396 -----IQMLDQRTSLSTELERCKVEEESKSDMESLTLAL-----QEASTESSA 441  
OY 335 KATLKTNNKTKGADLEKLFESVKNLSKAKE-----MLTNSVKEITLS 378  
DB 442 KATLVQCEELKNC-----ESQVDSLKLASKETNKEKEMLEDANNEIDS 486

RESULT 7  
US-09-708-427-15045  
; Sequence 15045, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15045  
; LENGTH: 1304  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1304  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1304  
; OTHER INFORMATION: Ceres Seq. ID 1828628  
US-09-708-427-15045

Query Match  
Best Local Similarity 9.18; Score 170; DB 6; Length 1304;  
Matches 112; Conservative 60; Mismatches 136; Indels 204; Gaps 21;  
QY 24 PNLTEINKKIDTSNAV-----LLAVKEVALISIDETAAKAIKGIHONGDPTENN 76  
DB 13 PRLSKLSKSDSASAPKVASRLVKTGELQTLQIQEDLKR-----DEQ-- 59  
QY 77 HNGSLGAVAIPTLIKOR---LDGLK-----NEGLEKIDAKKCSFT 118  
DB 60 -----IELKKDKAKAIDDLKESKLEVEANKEKLAQAESFEVEKFR 108  
QY 119 -----TNK-----LKEKHT-DLG-----KEGYTDADAKA 142  
DB 109 AVELEQNGLEAVOKKDVTSKNELESIRSOHALDISALSTTELOVRKHELSMTADANK 168  
QY 143 ILKANGTKTKGAE-----ELGKL-----FESVEVLSKAKE----- 173  
DB 169 ALSHAEARIRAEIHAKEAELILASELGRKLALGSKKEKEKALEGNEIVSKLSETELLRG 228  
QY 174 -----MLANSYKE-----LTSVVAESPCKPSMNNSGKGNIS 207  
DB 229 ELEKVSILSSSLKEDEGLVEQLKVDLEAKMAESCSTNSVEEMKMKVHELKEVEESNRS 288  
QY 208 ANSADESVKGMPLTETSKKITESNAV-----LAVKEVETLLTIDE--- 249  
DB 289 KSSASESME-----SYMQLAELNHVLEHTSDNAAKKEKTELLEKTEAQRTDLEEYGR 343  
QY 250 -----LKAIGKTKIKNV-----SLDNEADHNGSLISGAYLISNLTIRKIS 290  
DB 344 QVCIAKEKSKLENLVESIKSELETSOBEKTRALDNEKAATSN-----IQNLDDRT 396  
QY 291 AIKDSGELKAELIEKAKCSEFTAKLGEHTDLKEGYTDNAKALILKNNDKTGAD 350  
DB 397 LSTELERCKVEEKSCKMESTIAL-----QEASTSESAKATLVCOEELKNC-- 446  
QY 351 LEKLFESYKNISSKAKE-----MLTNSVKELTNS 378  
DB 447 -ESQVDSIKLASKETNEKYMEDARNIDS 477

APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19883  
; LENGTH: 1014  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1014  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1014  
; OTHER INFORMATION: Ceres Seq. ID 1836847  
US-09-708-427-19883

Query Match  
Best Local Similarity 9.08; Score 167.5; DB 6; Length 1014;  
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

QY 8 KDGNANSASDE-----SVKGPNL-----TEINKKIDTS-NAVLAWE 45  
DB 401 KHGEADSKGYLGQVAVELSTLEAFYKSSLSLEALNITENEKELTENLVNVSSEKKR 460  
QY 46 VEALLSIDETIAK-----AIGKIHONGDPTENNNGSLAGAYASTL--IKO 94  
DB 461 LEA---TVDESVKISSESENLIESIRNELNVTOG-KLESTENDLKAAGQESVEVEXLS 516  
QY 95 KLDGKNEGLEKEKIDAA-----KKCSEFTNK-----LKEKH 126  
DB 517 AEESELEQKG--REIDETITRMELALHOSLSDSEHRLQKAMEFTSDESSALTEKL 574  
QY 127 TDL-GKEGYTDADAKAEILKANGTKGAEELGKLFESVEVLSKAKEMLANSVREL--- 182  
DB 575 RDLBGKIKSYEQLAASGSSSLKEKLEQTLGRLAASVNEKIKQEPDQAQESLOSS 634  
QY 183 -TSPVVAESPCK-----PSMVNNSGKDGNISANSDESCKGNLIEISKKITESNAV 233  
DB 635 SESELLAETNNQKIKIQELEGIGSGVEKETALKRLEALE-----RFQKETEESDL 689  
QY 234 VLAKEVETLLISIDELA-KAIG-----KKIKN----- 260  
DB 690 VEKLTGHENQIEEYKRLAEMASGVADTRKVELEDALSKLKNLESTIEELGAKCOGLEKES 749  
QY 261 -----DVSLDNEADHNGSLISGAYLISNLTIRKISAIKDSGELKA-EIEKAKKCSSEFT 313  
DB 750 GDIAEVNKLNLNLNHS-----ENAELOTKLSALEAKEDQYAEVLEASKTTIIDLT 802  
QY 314 AKL--KGEHTDLKGEYTDNAKAILKTNNDKTGGADELEKLEESYK-NLSKAKEMLT 370  
DB 803 KQLTSEGEKLSQLEKRAVAEAKSVLESH-----FEELKELTSEVKAOLKEVENAAT 856  
QY 371 NSVK--ELTS 378  
DB 857 ASVKAELTNS 866

RESULT 9  
US-09-708-427-19882  
; Sequence 19882, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09

NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19882  
LENGTH: 1018  
TYPE: PRF  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..1018  
OTHER INFORMATION: Xaa is any amino acid  
LOCATION: 1..1018  
OTHER INFORMATION: Ceres Seq. ID 1836846  
US-09-708-427-19882

Query Match 9.0%; Score 167.5; DB 6; Length 1018;  
Best Local Similarity 23.1%; Pred. No. 0.002;  
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

QY 8 KDGNTSANSAD-----SVKGNL-----TEINKKITDS-NAVLLAVKE 45  
QY 405 KHGETEADSKGYLQVAVELSTLEAFQVKSLSLEALNINATENKEKLETNINAVTSEKK 464  
QY 46 VEALLSIDETIAK-----AIGKKIHONNGLOTEENNNGSLAGAVAI STL--IKQ 94  
QY 465 LEA---TVDEYSVKISESENLESI RNELNVTQG-KLESLENDLKAAGLOESEYMERLKS 520  
QY 95 KLGLKNEGLKEKIDAA-----KKCSEFTFNK-----LKEKH 126  
QY 521 AESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLOKAMEFTSRSEASLSLEKL 578  
QY 127 TDL-GKEGYTDADAKAELKANGTKTKGAELGKLFESVEYLSKAEMLANSVKEL--- 182  
QY 579 RDEGKIKSYEEQVLAESGKSSLEKELEQTLGRLLAAESVNEKLEQFDQAQEKSLQSS 638  
QY 183 -TSPVVAESPK-----PSMVNNGSGKGNTSANSADSVKGNLTKTSKTTESNAV 233  
QY 639 SESELLAETNNQIKIKIOLEGLIGSGVEKETALKRLEAIE---RFNOKETESSDL 693  
QY 234 VLAWEVETLLTSIDELA-KAIG-----KKIKN----- 260  
QY 694 VEKLTHTENQIEEYKKAHLAHSAGVADRKYVELEDALSKLNLESTIEELKACGGLLEKS 753  
QY 261 -----DVSIDNEADHNGSLISGAYLISNLTTKKISAIKDSGELKA-EIEKAKKCSSEFT 313  
QY 754 GDLEAVNLKLNLELANHGS-----EANELQTKLSALEAKEQDTANELKESKTTIEDLT 806  
QY 314 AKL--KGEHTDLCGEVTDNNAKKAILEKTNNDKTKGADLEKLEFSYK-NLSKAKKELT 370  
QY 807 KQLTSEGEKLSQIEKRLAFAAERSVLESH-----FEELKTLSEVKAQOLKEVNAENAT 860  
QY 371 NSVK--ELTS 378  
QY 861 ASYKVAELTS 870  
DB

RESULT 10  
US-09-708-427-19881  
Sequence 19881, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19881  
LENGTH: 1269  
TYPE: PRF  
ORGANISM: Arabidopsis thaliana

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..1269  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..1269  
OTHER INFORMATION: Ceres Seq. ID 1836845  
US-09-708-427-19881

Query Match 9.0%; Score 167.5; DB 6; Length 1269;  
Best Local Similarity 23.1%; Pred. No. 0.0027;  
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

QY 8 KDGNTSANSAD-----SVKGNL-----TEINKKITDS-NAVLLAVKE 45  
QY 656 KHGETEADSKGYLQVAVELSTLEAFQVKSLSLEALNINATENKEKLETNINAVTSEKK 715  
QY 46 VEALLSIDETIAK-----AIGKKIHONNGLOTEENNNGSLAGAVAI STL--IKQ 94  
QY 716 LEA---TVDEYSVKISESENLESI RNELNVTQG-KLESLENDLKAAGLOESEYMERLKS 771  
QY 95 KLGLKNEGLKEKIDAA-----KKCSEFTFNK-----LKEKH 126  
QY 772 AESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLOKAMEFTSRSEASLSLEKL 829  
QY 127 TDL-GKEGYTDADAKAELKANGTKTKGAELGKLFESVEYLSKAEMLANSVKEL--- 182  
QY 830 RDEGKIKSYEEQVLAESGKSSLEKELEQTLGRLLAAESVNEKLEQFDQAQEKSLQSS 889  
QY 183 -TSPVVAESPK-----PSMVNNGSGKGNTSANSADSVKGNLTKTSKTTESNAV 233  
QY 890 SESELLAETNNQIKIKIOLEGLIGSGVEKETALKRLEAIE---RFNOKETESSDL 944  
QY 234 VLAWEVETLLTSIDELA-KAIG-----KKIKN----- 260  
QY 945 VEKLTHTENQIEEYKKAHLAHSAGVADRKYVELEDALSKLNLESTIEELKACGGLLEKS 1004  
QY 261 -----DVSIDNEADHNGSLISGAYLISNLTTKKISAIKDSGELKA-EIEKAKKCSSEFT 313  
QY 1005 GDLEAVNLKLNLELANHGS-----EANELQTKLSALEAKEQDTANELKESKTTIEDLT 1057  
QY 314 AKL--KGEHTDLCGEVTDNNAKKAILEKTNNDKTKGADLEKLEFSYK-NLSKAKKELT 370  
QY 1058 KQLTSEGEKLSQIEKRLAFAAERSVLESH-----FEELKTLSEVKAQOLKEVNAENAT 1111  
QY 371 NSVK--ELTS 378  
QY 1112 ASYKVAELTS 1121  
DB

RESULT 11  
US-09-815-242-5883  
Sequence 5883, Application US/09815242  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727

;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5883  
;; LENGTH: 837  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-815-242-5883

Query Match 8.7%; Score 162.5; DB 6; Length 837;  
Best Local Similarity 18.7%; Pred. No. 0.0033;  
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSEYKGP--NLTETINK-----KITDSNAVLAVKEVEALLSS 52  
DB 350 NGNQVYANAKTTAKNALNLTSTINNAQKEALKSQIEGATTVAGVNOVSTTASELNTAMSN 409  
QY 53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGLAGAVAISTLLKQKLDGLKNGEKER 107  
DB 410 LQNGINDEATKA--ALNGTONLEKAKQHANTAIIDL-----SHLTNAQKEALKQ 458  
QY 108 IDAAKCEFTFTNKLKEKHTDGLKEGVTDADAKAELILKANGTKTKGAELKLFESVEVL 167  
DB 459 VOOSTTVAAGQNEQKANNVDAAMDKLROSIDNATTKONONYTDSQKKKAYNNAYTT 518  
QY 168 SRAAKEMLANSVKELTSPVVAESPCKPSKVNNSGKDNSTANSADSEYKGPNTETISKI 227  
DB 519 AGCI-----IDQTSPL-----DPTVINOAGQVSTTKNALN--ENLEAKQQA 563  
QY 228 TESNAVLAVKEVEELLTSIDELAKAIGKKIKNDVSLDNEADHNSLISGAVLISNLT 287  
DB 564 SOS-----LGLDLNNAQKQTVTDQINGAHYVDEANQIKONQNLNTAMGN 610  
QY 288 KISAIDSELSKAEI-----EKAKK-----CSEFTAKLKGEHTDGL--KEGVTDN 332  
DB 611 LKQALADKATKATVFTDADQAKQAVNTAVTNMENIISKANGNATQAEVEQAIKQVN 670  
DB 333 AKKALITNNDKTKGADELEKLFESYKULSKAAKEMLTNSVELTS 378  
DB 671 AAKQALNGNANVQAKDEATALINSNDLNAQKQALKQVONAYTT 716

RESULT 12  
US-10-072-851-5883  
; Sequence 5883, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Yamamoto, John D.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits  
; FILE REFERENCE: ELITRA.028A  
; CURRENT APPLICATION NUMBER: US/10/072,851

;; CURRENT FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: 60/267,636  
;; PRIOR FILING DATE: 2001-02-09  
;; NUMBER OF SEQ ID NOS: 15811  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5883  
;; LENGTH: 837  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-10-072-851-5883

Query Match 8.7%; Score 162.5; DB 7; Length 837;  
Best Local Similarity 18.7%; Pred. No. 0.0033;  
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSEYKGP--NLTETINK-----KITDSNAVLAVKEVEALLSS 52  
DB 350 NGNQVYANAKTTAKNALNLTSTINNAQKEALKSQIEGATTVAGVNOVSTTASELNTAMSN 409  
QY 53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGLAGAVAISTLLKQKLDGLKNGEKER 107  
DB 410 LQNGINDEATKA--ALNGTONLEKAKQHANTAIIDL-----SHLTNAQKEALKQ 458  
QY 108 IDAAKCEFTFTNKLKEKHTDGLKEGVTDADAKAELILKANGTKTKGAELKLFESVEVL 167  
DB 459 VOOSTTVAAGQNEQKANNVDAAMDKLROSIDNATTKONONYTDSQKKKAYNNAYTT 518  
QY 168 SRAAKEMLANSVKELTSPVVAESPCKPSKVNNSGKDNSTANSADSEYKGPNTETISKI 227  
DB 519 AGCI-----IDQTSPL-----DPTVINOAGQVSTTKNALN--ENLEAKQQA 563  
QY 228 TESNAVLAVKEVEELLTSIDELAKAIGKKIKNDVSLDNEADHNSLISGAVLISNLT 287  
DB 564 SOS-----LGLDLNNAQKQTVTDQINGAHYVDEANQIKONQNLNTAMGN 610  
QY 288 KISAIDSELSKAEI-----EKAKK-----CSEFTAKLKGEHTDGL--KEGVTDN 332  
DB 611 LKQALADKATKATVFTDADQAKQAVNTAVTNMENIISKANGNATQAEVEQAIKQVN 670  
QY 333 AKKALITNNDKTKGADELEKLFESYKULSKAAKEMLTNSVELTS 378  
DB 671 AAKQALNGNANVQAKDEATALINSNDLNAQKQALKQVONAYTT 716

RESULT 13  
US-09-815-242-13080  
; Sequence 13080, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931



;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 13080  
;; LENGTH: 875  
;; TYPE: PRF  
;; ORGANISM: Staphylococcus aureus  
US-09-815-242-13080

Query Match 8.7%; Score 162.5; DB 6; Length 875;  
Best Local Similarity 18.7%; Pred. No. 0.0035;  
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSVKGP--NLTEINK-----KITDSNAVLAVKEVEALLSS 52  
DB 387 NGNONVANAKTTAKNMLNLTSTINNOKKALKSQIEGATYAGVNOYSTASELNTAMSN 446  
QY 53 I-----DEIAKAIGKKIHONNGIDTENNHGSLLAGAVASTLIKOKLDGKNEGLKEK 107  
DB 447 LONGINDEAATKA---ALNGTONLEKAKOHANTAIDGL-----SHLTWAQKEALKOL 495  
QY 108 IDAKKSEFTTNKLEKHHDGEGVTDADAKAAILKANGTKTKGAELGKLEESYEV 167  
DB 496 VQOSTTVAEAGNEOKANNVDAMDKROSTADNATTKONQNTDASONKKDAYNNAVTT 555  
QY 168 SKAKKEMIANSVKELTSPVVAESPCKPSMVNSGDKGNTSANSADSVKGPNTLEISKI 227  
DB 556 AAGI-----IDGTTSPTL-----DPTVINQAAQGVSTTKNALNGN---ENLEAAQQA 600  
QY 228 TESNAVVLAVKEVEETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLT 287  
DB 601 SOS-----LGSIDLNNNAOKQVTTDQINGAHVYDEANOIKONQONNTAMGN 647  
QY 288 KISAIKDSGELKAEI-----EKAKK-----CSEFTAKLKEHHDGLG--KEGYDDN 332  
DB 648 LKQAIADKATKATVNTDADQAKQOAVNTAVTNAENITISANGNATQAEVEAIOKOVN 707  
QY 333 AKKAILTKNDKTGADELEKLFESVKNLSKRAKEMLTNSYKELTS 378  
DB 708 AAKQALNGNANVQAKDEATLINSNDLNOAQDALKQOVQNTT 753

RESULT 14  
US-10-072-851-13080  
;; Sequence 13080, Application US/10072851  
;; GENERAL INFORMATION:  
;; APPLICANT: Carr, Grant J.  
;; APPLICANT: Xu, H. Howard  
;; APPLICANT: Foulkes, J. Gordon  
;; APPLICANT: Zamudio, Carlos  
;; APPLICANT: Haselbeck, Robert  
;; APPLICANT: Ohlsen, Karl L.  
;; APPLICANT: Zyskind, Judith W.  
;; APPLICANT: Wall, Daniel  
;; APPLICANT: Trawick, John D.  
;; APPLICANT: Yamamoto, Robert T.  
;; APPLICANT: Roemer, Terry  
;; APPLICANT: Jiang, Bo  
;; APPLICANT: Boone, Charles  
;; APPLICANT: Bussey, Howard  
;; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits  
;; FILE REFERENCE: ELITRA.028A  
;; CURRENT APPLICATION NUMBER: US/10/072,851  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: 60/267,636  
;; PRIOR FILING DATE: 2001-02-09  
;; NUMBER OF SEQ ID NOS: 15811  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 13080

;; LENGTH: 875  
;; TYPE: PRF  
;; ORGANISM: Staphylococcus aureus  
US-10-072-851-13080

Query Match 8.7%; Score 162.5; DB 7; Length 875;  
Best Local Similarity 18.7%; Pred. No. 0.0035;  
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSVKGP--NLTEINK-----KITDSNAVLAVKEVEALLSS 52  
DB 387 NGNONVANAKTTAKNMLNLTSTINNOKKALKSQIEGATYAGVNOYSTASELNTAMSN 446  
QY 53 I-----DEIAKAIGKKIHONNGIDTENNHGSLLAGAVASTLIKOKLDGKNEGLKEK 107  
DB 447 LONGINDEAATKA---ALNGTONLEKAKOHANTAIDGL-----SHLTWAQKEALKOL 495  
QY 108 IDAKKSEFTTNKLEKHHDGEGVTDADAKAAILKANGTKTKGAELGKLEESYEV 167  
DB 496 VQOSTTVAEAGNEOKANNVDAMDKROSTADNATTKONQNTDASONKKDAYNNAVTT 555  
QY 168 SKAKKEMIANSVKELTSPVVAESPCKPSMVNSGDKGNTSANSADSVKGPNTLEISKI 227  
DB 556 AAGI-----IDGTTSPTL-----DPTVINQAAQGVSTTKNALNGN---ENLEAAQQA 600  
QY 228 TESNAVVLAVKEVEETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLT 287  
DB 601 SOS-----LGSIDLNNNAOKQVTTDQINGAHVYDEANOIKONQONNTAMGN 647  
QY 288 KISAIKDSGELKAEI-----EKAKK-----CSEFTAKLKEHHDGLG--KEGYDDN 332  
DB 648 LKQAIADKATKATVNTDADQAKQOAVNTAVTNAENITISANGNATQAEVEAIOKOVN 707  
QY 333 AKKAILTKNDKTGADELEKLFESVKNLSKRAKEMLTNSYKELTS 378  
DB 708 AAKQALNGNANVQAKDEATLINSNDLNOAQDALKQOVQNTT 753

RESULT 15  
US-09-815-242-5835  
;; Sequence 5835, Application US/09815242  
;; GENERAL INFORMATION:  
;; APPLICANT: Haselbeck, Robert  
;; APPLICANT: Ohlsen, Karl L.  
;; APPLICANT: Zyskind, Judith W.  
;; APPLICANT: Wall, Daniel  
;; APPLICANT: Trawick, John D.  
;; APPLICANT: Carr, Grant J.  
;; APPLICANT: Yamamoto, Robert T.  
;; APPLICANT: Xu, H. Howard  
;; TITLE OF INVENTION: Identification of Essential Genes in  
;; FILE REFERENCE: ELITRA.011A  
;; CURRENT APPLICATION NUMBER: US/09/815,242  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: 60/191,078  
;; PRIOR FILING DATE: 2000-03-21  
;; PRIOR APPLICATION NUMBER: 60/206,848  
;; PRIOR FILING DATE: 2000-05-23  
;; PRIOR APPLICATION NUMBER: 60/207,727  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5835



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:42 ; Search time 39.62 Seconds

(without alignments)  
349.806 Million cell updates/sec

Title: US-09-596-746A-28

Sequence: 1859  
1 MACNNSGKDGNTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Partial number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database : SWISSPROT\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	50.8	210	1	OSCL_BORBU
2	631	33.9	212	1	OSCL_BORBU
3	397.5	21.4	214	1	VM24_BORHE
4	397	21.4	215	1	VM03_BORHE
5	179	9.6	1509	1	MYSN_ACACA
6	173.5	9.3	1957	1	YD86_SCHPO
7	168.5	9.1	483	1	M6_STRPY
8	168.5	9.1	564	1	M12_STRPY
9	167.5	9.0	998	1	SCA4_RICAF
10	166.5	9.0	1251	1	RBP2_PLAVB
11	165.5	8.9	1938	1	MYSL_AEQIR
12	164	8.8	1790	1	USO1_YEAST
13	163.5	8.8	2116	1	MYSL_DICDI
14	163	8.8	3210	1	CENF_HUMAN
15	157.5	8.5	1966	1	MYSL_CAEEL
16	157	8.4	1786	1	LM1L_MOUSE
17	155.5	8.4	492	1	MS_STRPY
18	154	8.3	1423	1	SCA4_RICRH
19	153.5	8.3	1017	1	REST_HUMAN
20	153.5	8.3	2869	1	RBP1_PLAVB
21	149.5	8.0	1935	1	MYSL_CYPCA
22	149	8.0	1011	1	SCA4_RICAF
23	148.5	8.0	1189	1	SC11_CHICK
24	148	8.0	866	1	MYSP_SCHJA
25	148	8.0	875	1	Z1PL_YEAST
26	148	8.0	2017	1	MYSL_DROME
27	148	8.0	3110	1	LM2_HUMAN
28	147.5	7.9	705	1	YNP9_CAEEL
29	147.5	7.9	1639	1	LMG1_DROME
30	147	7.9	991	1	SCA4_RICAU
31	147	7.9	1500	1	SPS5_STRGN
32	147	7.9	1961	1	MYH9_RAT
33	146.5	7.9	967	1	KINH_LOLPE

34	146	7.9	679	1	YK99_YEAST
35	145.5	7.8	1786	1	LM1L_HUMAN
36	145	7.8	866	1	MYSP_SCHMA
37	145	7.8	1010	1	SCA4_RICPA
38	144.5	7.8	1164	1	BAG_STRAG
39	144.5	7.8	1433	1	REST_CHICK
40	144.5	7.8	1453	1	Y373_BOVIN
41	144.5	7.8	2349	1	TPR_HUMAN
42	144	7.7	1713	1	LM2_HUMAN
43	143.5	7.7	1935	1	MYH7_HUMAN
44	143	7.7	775	1	YHGE_BACSU
45	143	7.7	3672	1	LM2_CAEEL

## ALIGNMENTS

RESULT	1	STANDARD	PRT	210 AA.
OSCL_BORBU	007337			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB819.			
OG	Borrelia burgdorferi (Lyme disease spirochete).			
OS	Plasmid lp54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=93268136; PubMed=8098841;			
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,			
RA	Will G., Wilske B.;			
RT	"Genetic heterogeneity of the genes coding for the outer surface			
RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";			
RL	Med. Microbiol. Immunol. 182:37-50(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=9329332; PubMed=8478108;			
RA	Wilske B., Preac-Mursic V., Jauris S., Soutschek E.,			
RA	Schwab E., Wanner G.;			
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant			
RT	major outer surface protein of Borrelia burgdorferi.";			
RL	Infect. Immun. 61:2182-2191(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=94041630; PubMed=8225587;			
RA	Pedula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;			
RT	"Molecular characterization and expression of p23 (OspC) from a North			
RT	American strain of Borrelia burgdorferi.";			
RL	Infect. Immun. 61:5097-5105(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=96025162; PubMed=7494039;			
RA	Eukunaga M., Hamase A.;			
RT	"Outer surface protein C gene sequence analysis of Borrelia			
RT	burgdorferi sensu lato isolates from Japan.";			
RL	J. Clin. Microbiol. 33:2415-2420(1995).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=98065943; PubMed=9403685;			
RA	Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,			
RA	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			

```

RA RA Soutschek E.;
RT "Molecular analysis and expression of a Borrelia burgdorferi gene
RL encoding a 22 kDa protein (pc) in Escherichia coli.";
RM Mol. Microbiol. 6:503-509(1992).
RN [2]
RP SEQUENCE OF 1-205 FROM N.A.
RR STRAIN-DK26;
RX MEDLINE=94075528; PubMed=8253951;
RA Thiesen M., Frederiksen B., Lebesch A.M., Vuust J., Hansen K.;
RT "Polymorphism in ospC gene of Borrelia burgdorferi and
RL immunoreactivity of ospC protein: implications for taxonomy and for
RM use of ospC protein as a diagnostic antigen.";
RN J. Clin. Microbiol. 31:2570-2576(1993).
CC -I- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to licenses@isb-sdb.ch).
CC -----
DR EMBL: X62169; CAA4093.1; -
DR DR EMBL: X73624; CAA52003.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW Outer membrane; lipoprotein; signal; plasmid; antigen.
FT SIGNAL 1 18 BY SIMILARITY
FT CHAIN 19 212 OUTER SURFACE PROTEIN C.
FT FT 19 N-ACTYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE 212 AA; 22499 MW; C206C231EFB2ETD4 CRC64;

Query Match 33.9%; Score 631; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. NO. 2,6e-22;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNNSGKRGN-SANSDESVKGNTLEINKITDSNAVLLAVEYALLSIDETAAR 59
DB :|||||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 IECNNNGGGSDASTNPADESAAGPNELEISKITDSNAFLAVKEVETLVLSIDEIAKK 76
QY 60 AIGRKIHONNGLDTEENNNGSLLAGAVALSTLIKLOKLDGLKN-EGLKEKIDAARKSETF 118
DB 77 AIGQRKDINNNGALANNONGSLLAGAVAIISPLTEIKSKLNLEELKETLTAAKAKCEEF 136
QY 119 TTKLEKERTDGRKGVTPADAKKAIVIKANGKKTGADELGLPFSSVYLSKAEMIAN 178
DB :||||:||||: |||||:||||:||||:||||:||||:||||:||||:
137 TNKLKSGHADLGKODATVDHAKAAILTHATTDDGAEKFEDFESVGILKAQVAOVLTS 196
QY 179 VKELTSPVAESPCK 194
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 VKELTSPVAESPCK 212

RESULT 3
VM24_BORHE ID VM24_BORHE STANDARD; PRT; 214 AA.
AC P32778;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN VMP24.
OS Borrelia hermsli.
OC Plasmid.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN-SSP. HSI SEROTYPE 24;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbours A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids
are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04786; AAA22964.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
FT SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;
SO

```

```

Query Match 21.4%; Score 397.5; DB 1; Length 214;
Best Local Similarity 44.8%; Pred. No. 6.7e-12;
Matches 91; Conservative 37; Mismatches 58; Indels 17; Gaps 5;

```

```

OY 1 MACNSGKCGNTSANSADSVKGP----NLTEINKKITDSNAVLAVKEVALLSIDET 56
DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKVSKRIKASAPASVKEVETLVKVSDEL 71
OY 57 AKAIGKTIHONNGLDTEENHNSLAGAVASTLIKOKL-----DGLKNEGLKKEIDA 110
DB 72 -AKAIGKTIHONNGLDTEENHNSLAGAVASTLIKOKL-----DGLKNEGLKKEIDA 110
OY 111 AKKSEFTFNKLEKHTDLEGVTDADAKKALIKANGTKGAEELKLFESVEVLSKA 170
DB 130 VKSKAEAFPLNKLKDGHTLKGKNDSDDTKKAIKKNSDKTKGASELEALNTAVDALLKA 189
OY 171 AKEMLANSVKELTSPVVAESPKK 193
DB 190 AEGVEAAIKELTAPVKAEPKPSQ 212

```

```

RESULT 4
VMP3_BORNE
ID VMP3_BORNE STANDARD; PRT; 215 AA.
AC 002448;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN VMP3.
OS Borrelia hermsli.
OC Plasmid.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_Taxid=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SSP. HSI SEROTYPE 3;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbours A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

```

```

RT are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04789; AAA22967.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
FT SEQUENCE 215 AA; 23139 MW; 684C74D35F87C771 CRC64;
SO

```

```

Query Match 21.4%; Score 397; DB 1; Length 215;
Best Local Similarity 45.3%; Pred. No. 7.1e-12;
Matches 92; Conservative 35; Mismatches 60; Indels 16; Gaps 5;

```

```

OY 1 MACNSGKCGNTSANSADSVKGP----NLTEINKKITDSNAVLAVKEVALLSIDET 56
DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKVSKRIKASAPASVKEVETLVKVSDEL 71
OY 57 AKAIGKTIHONNGLDTEENHNSLAGAVASTLIKOKL-----DGLKNEGLKKEIDA 110
DB 72 -AKAIGKTIHONNGLDTEENHNSLAGAVASTLIKOKL-----DGLKNEGLKKEIDA 110
OY 111 AKKSEFTFNKLEKHTDLEGVTDADAKKALIKANGTKGAEELKLFESVEVLSKA 170
DB 131 VKKSEAFYTVQVSKHTDLAKKEGVTDAHAKSALILVDTGDKGAALIKLNTAIDELKA 190
OY 171 AKEMLANSVKELTSPVVAESPKK 193
DB 191 ANDAVETVIRKELTASVKAEPKPSQ 213

```

```

RESULT 5
MYSN_ACACA
ID MYSN_ACACA STANDARD; PRT; 1509 AA.
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_Taxid=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87308395; PubMed-3040773;
RA Hammer J.A. II, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail."
RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

```

CC LIGHT CHAIN SUBUNITS (MUC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MC-2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS  
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE  
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.  
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING  
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y00624; CAA68663.1; -  
 DR PIR: A27224; A27224.  
 DR HSSP: P08799; LMND.  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IO; 2.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRODOM: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IO; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IO; 1.  
 DR Myosin: Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;  
 KM Methylation; Alkylation; Phosphorylation; Multigene family.  
 FT DOMAIN 1 789  
 FT MYOSIN HEAD-LIKE.  
 FT 790 819  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 848 1509  
 FT ALPHA-HELICAL TAILPIECE (S2).  
 FT DOMAIN 1227 1252  
 FT HINGE.  
 FT DOMAIN 1253 1509  
 FT LIGHT MEROMYOSIN (LMN).  
 FT DOMAIN 1253 1482  
 FT ALPHA-HELICAL TAILPIECE (LMN).  
 FT NP\_BIND 1483 1509  
 FT 182 189  
 FT NONHELICAL TAILPIECE.  
 FT DOMAIN 660 682  
 FT ACTIN-BINDING.  
 FT DOMAIN 766 780  
 FT ACTIN-BINDING.  
 FT MOD\_RES 133 133  
 FT METHYLATION (TRI-) (POTENTIAL).  
 FT MOD\_RES 700 700  
 FT ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 1489 1489  
 FT PHOSPHORYLATION.  
 FT MOD\_RES 1494 1494  
 FT PHOSPHORYLATION.  
 FT MOD\_RES 1499 1499  
 FT PHOSPHORYLATION.  
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 9.6%; Score 179; DB 1; Length 1509;  
 Best Local Similarity 24.9%; Pred. No. 0.33;  
 Matches 102; Conservative 69; Mismatches 146; Indels 92; Gaps 20;

QY 28 EINKKITSNVLAVKVEEA---LSSIDEIAKAIGKIHQ--NNGLDTENNHNSGL 81  
 DB 934 ELQETSASNDLEQKRLTEAKGELKAKSLSE---EENKRAALQEAATKVESENNELQDK 990  
 QY 82 LAGAAVAITLKQKLDGKLNGLKEKIDA---AKKSEFTFNKLEKHTDIEGKGVYDA- 137  
 DB 991 YEDAAHAHDSKKKEEDLSRE-LRETKDALADANISSETLSKSL- KNTERGADVANEL 1047  
 QY 138 -DAEALIKANGTGTGAEEGLKEEVEYLSKAKEMLSNVELTSPV----- 186  
 DB 1048 DDVATATKQLEKTKKSLSEELAQTRADLEE-EKSGKEAASAKAQLOQOQEDARSEYDSL 1106  
 QY 187 -----VASPKPSMVNNSG-----KDGNTSANSADSVK--PNLTETSKKIT-----ES 230  
 DB 1107 KSKLSAEKSLKTAQDNRDLQEDLEDERTVIRANVVKOKKALAKLTLEDVATALDQK 1166

QY 231 NAVLAKVEYELLTSLDEL-----AKAIGKINDVSLDNEADHNSLISGAYL 280  
 DB 1167 NA---AAQAQKTLKTYDDEYTERKRLLEAEASAAKLEKERN--ALBEVAQ----- 1210  
 QY 281 ISNLTIRKTSIAIDSG-----BLKAEIEKAK--CSDEPFAKIKGEHTDQK 325  
 DB 1211 ---LADDAEDRSDQAQRRKLNTRISLQSELENAKPTGASSEYKKEGELERLEE 1266  
 QY 326 EGYTDMAKKAIIKTYNDKTKGADELEKLFESVKNISLRAKEMLTNSVK 374  
 DB 1267 ELTLAQEAARAAR-KNLKAN--LELEELRQEDADAARDNDKLVKDNK 1312

RESULT 6  
 ID YD86\_SCHPO STANDARD. PRT; 1957 AA.  
 AC 010411:  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL. 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME 1.  
 GN SPAC1F3.06C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 CC NCBI\_TaxID=4896;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=972;  
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z70690; CAA94624.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 9.3%; Score 173.5; DB 1; Length 1957;  
 Best Local Similarity 22.8%; Pred. No. 0.76;  
 Matches 108; Conservative 73; Mismatches 148; Indels 145; Gaps 18;

QY 8 KDGNTSANSADSVKGPNTLEIRKKTDSNAVLAV-KVEVALLSIDEI----- 56  
 DB 360 KDSRTINSQLEEMVVLKLSNRT-IHSQLTDAESKLSFQENKSLKSIDLEYONNLSSK 418  
 QY 57 -----AKAIGKRIHONNGLDTEN-----NHNSGL 82  
 DB 419 DKWAKOVSSQLEEARSSLAHATGKLAELINSEKDPQNKIKODEPTEQDLKACLNSSNEL 478  
 QY 83 AGAYALSTLKKQKLDGKLNGLKEKIDAACKSEFTFNKL-----KEKH----- 126  
 DB 479 KER---SALIDKKDDELNN--LREQIKQKRVSESTOSLSLOPDLINERKKHVEYSQ 533  
 QY 127 -----TDLKEGVTADAKAEALIKANGTKTGABELGLKFESEVLSKAA 171  
 DB 534 LNELKGELOTETSEINSEHLSQSLTIAAKKAAVATN-----NELSEKNSLOTLNAP 586  
 QY 172 KEMLANSKVELTSPVVAASPKPSMVNNSGKDGNTSANSADSVKGPNTLEIRKKTEN 231  
 DB 587 QEKLAQSVQWL-----KENQNFSSLDTSFKKLINESHOELEN-----NQDTTKQKLDTS 636  
 QY 232 AVLVLAV-----EVELLTSLDELAKAIGKIKINDVSLDNEADHNS 273  
 DB 637 SKLOLOLERANFQKESTLSDENNDLRTKLKLEESNKSILRKQEDVSL----- 688

OY 274 LISGAVLISNLTKKISAIKDSGLKAEIEKAKKCESEFTAKLGEHTDGRK-----G 327.  
 DB 689 -----NKIOTLKEDLRKSEBALRFSKLEAKNL-REVYIDNKGHEHLEAGRNDLHSS 739  
 OY 328 VTDDNKKAKLKKNNNDKTKGADELKLFESVKNL---SKAKEMLTNSVKEITS 378  
 DB 740 LSDAKRTNML---SSELTKSSEDEVKRLTANVFETLTODSKAKMKSFTSLVNSYOS 791

RESULT 7  
 M6\_STRPY ID M6\_STRPY STANDARD: PRT: 483 AA.  
 AC P08089;  
 DT 01-AUG-1988 (rel. 08, Created)  
 DT 01-AUG-1988 (rel. 08, Last sequence update)  
 DT 20-AUG-2001 (rel. 40, Last annotation update)  
 DE M PROTEIN, SEROTYPE 6 PRECURSOR.  
 GN EMM6.  
 OS Streptococcus pyogenes.  
 NC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 NC Streptococcus.  
 NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86111835; PubMed=3511046;  
 RA Hollingshead S.K., Fischetti V.F., Scott J.R.;  
 RT "Complete nucleotide sequence of type 6 M protein of the group A  
 RT Streptococcus. Repetitive structure and membrane anchor."  
 RL J. Biol. Chem. 261:1677-1686(1986).  
 RN [2]  
 RP SEQUENCE OF 43-122 FROM N.A.  
 RX MEDLINE=85166224; PubMed=3885219;  
 RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;  
 RT "Relationship of M protein genes in group A streptococci."  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).  
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES  
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF  
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO  
 CC PHAGOCYTOSIS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.  
 CC -1- SIMILARITY: TO OTHER M PROTEINS.  
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS  
 CC IN THE REGION OF THE MEMBRANE ANCHOR.  
 CC -1- THE SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: M1338; AAA26920.1; -  
 CC PIR: A26297; A26297.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR003345; M\_repeat.  
 DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam: PF02370; M; 9.  
 DR PRINTS: PR00015; GP0SANCHOR.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
 KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
 KW Transmembrane; Coll; Signal.  
 FT SIGNAL 1 42  
 FT CHAIN 43 483 M PROTEIN, SEROTYPE 6.  
 FT DOMAIN 43 457 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 458 477 MEMBRANE ANCHOR.  
 FT DOMAIN 478 483 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 69 138 10 X 7 AA TANDEM REPEATS.  
 FT DOMAIN 157 347 4.5 X 25 AA TANDEM REPEATS.  
 FT DOMAIN 279 347 TWO DIRECTLY REPEATED 27 AMINO ACID  
 FT DOMAIN 348 411 BLOCKS SEPARATED BY 15 AMINO ACIDS.  
 FT HYDROPHILIC.

FT DOMAIN 412 448 GLY/PRO-RICH (CELL WALL-SPANNING).  
 FT DOMAIN 449 454 CONSERVED IN GRAM-POSITIVE COCCI SURFACE  
 FT PROTEINS.  
 SQ SEQUENCE 483 AA; 53472 MW; 68f87f28db53448 CRC64;  
 Query Match 9.1%; Score 168.5; DB 1; Length 483;  
 Best Local Similarity 22.9%; Pred. No. 0.27;  
 Matches 91; Conservative 57; Mismatches 153; Indels 97; Gaps 15;  
 OY 5 NSGRDNTSANSADSEYKGNLTETINKKIDTSDNNAVLAIVEEVALSSIDEIAKAKGK 64  
 DB 73 NNDKLTENNNDLNDQ---KNLTENKNLTDQKNLTENK-----NL 112  
 OY 65 IHONNGIDTENNHNHNSLAGAVAISTLIKOKLDGKNEGKEKIDAKKSEFTTNLKE 124  
 DB 113 TDONKNMTTENK---ELKAEENRLTT-----ENKGLTKKLEAE---EEANKERE 157  
 OY 125 KHTDLCK-EGVTDADAKKALILKANGTKTGAELGKFESEV---VLSKAKEM----- 174  
 DB 158 NKEAIGTLKTLDETIVYDKIAKQESKFT---ETIGLTKTLDETIVYDKIAKQESKFTTG 213  
 OY 175 -LANSVKEITSYVAESPKKPPSWNNSGKDNTSANSADSEYKGNLTETISKTITESNAV 233  
 DB 214 TLAKTLDETIVYDKIAKQESKFT-----GTLKTLDETIVYDKIAKQESKFTG 261  
 OY 234 VLAKEVEYELLTSLDELAKAIGKIKNDVSLDNEADHNSLISGAVLISNLTKKISAIK 293  
 DB 262 IGALKQELAKDKGNKVSASRKGLRDLASREA-----KKQVEK 302  
 OY 294 DSGELKAEIKAKKCESEFTAKLGEHTDGRKCEVYDDNKKALKTKNNNDKTKGADLEK 353  
 DB 303 DLNLTAELDKVKEEKOISDASRGRLRDASREAKKQVEKA-LEBANSKLALEKLNK 361  
 OY 354 LFESVKNL-----SKAKEMLTNSVKEIT 376  
 DB 362 ELESKTKTEKAEKLEAKLEAKLEAKLEAKLEAKLEAKLEAKLEAKLEAKLEAKLEAK 399

RESULT 8  
 M12\_STRPY ID M12\_STRPY STANDARD: PRT: 564 AA.  
 AC P19401;  
 DT 01-NOV-1990 (rel. 16, Created)  
 DT 01-NOV-1990 (rel. 16, Last sequence update)  
 DT 20-AUG-2001 (rel. 40, Last annotation update)  
 DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).  
 GN EMM12.  
 OS Streptococcus pyogenes.  
 NC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 NC Streptococcus.  
 NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX SRRAIN-CS24 / Serotype M12;  
 RX MEDLINE=88058777; PubMed=2445730;  
 RA Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;  
 RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream  
 RT sequences."  
 RL J. Bacteriol. 169:5633-5640(1987).  
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES  
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF  
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO  
 CC PHAGOCYTOSIS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.  
 CC -1- SIMILARITY: TO OTHER M PROTEINS.  
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS  
 CC IN THE REGION OF THE MEMBRANE ANCHOR.  
 CC -1- THE SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M18269; AAA88573.1; -  
 DR PIR: A60115; A60115  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR003345; M\_repeat.  
 DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam: PF02370; M; 9.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
 DR Virulence: Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
 KW Transmembrane; Coiled coil; Signal.  
 FT SIGNAL 1 41  
 FT CHAIN 42 >564 M PROTEIN, SEROTYPE 12.  
 FT DOMAIN 42 550 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 551 >564 MEMBRANE ANCHOR.  
 FT DOMAIN 44 505 COILED COIL (POTENTIAL).  
 FT DOMAIN 505 541 GUY/PRO-RICH (CELL WALL-SPANNING).  
 FT DOMAIN 542 547 CONSERVED IN GRAM-POSITIVE COCCI SURFACE  
 FT PROTEINS.  
 FT NON\_TER 564 564  
 FT SEQUENCE 564 AA; 62904 MW; 5F1549DACAA77B46 CRC64;

Query Match 9.1%; Score 168.5; DB 1; Length 564;  
 Best Local Similarity 22.9%; Pred. No. 0.32;  
 Matches 97; Conservative 59; Mismatches 166; Indels 101; Gaps 17;

QY 10 GWTMSADESVKGPNTLEINKKTTSDNAVLAV-REVEALLSSIDELAAKIGKTHON 68  
 DB 115 GLGIDNADLAK--ITELKSEVEKNVLSQIKKELEAKNDIQ-----FGEVHAA 165  
 QY 69 NGIDTENNHNGLSLAGAVAI STLKOKLDLGNKGLKEKIDAAKCEFTTNKLEKND 128  
 DB 166 DLT--RHKEIDAEKEVVIS--KLNG-ELDPLKOKVD-----ETDRNLQEKQKV 209  
 QY 129 LGKE--CVTDADAKKEAI-IRANGTKTGAEELGKLFEESEVLSRAK--EMIANSVKEL 182  
 DB 210 LSLDQOLAVTKENAKKDFELALGHOLADKEYNAKIAELSEKIDAKKPELALAGHQA 269  
 QY 183 TSPVVAESPKKPVMV-----NSGKDGNTSANSADSVKGPNTLEISK-----KI 227  
 DB 270 HNEQVALAEKDOIKOLEKOKIOLDSRKGTARDELAVROAKKATEALNLMKELKV 329  
 QY 228 TESNAVLAVK-----EVEITLTSTIDE--LAKAIKKIKNDVSLNDEA 268  
 DB 330 TEOKOILDSRKGTARDELAVRKSKQOYEALKOLEBQNKISEASRKGLRDLDTISREA 389  
 QY 269 DHNGSLISGAVYLSNLTIKKISAIKSGELAKAEIKAKKCESEFTAKLGEHTDLGREGV 328  
 DB 390 -----KQOVEDLANLTAEIDKVEKEKQISDASQGLRDLDSRE 430  
 QY 329 TDNNAKAILKTNDKTGADELEKLFESVKL-----SKAKEMLTNSV 373  
 DB 431 AKKQVEKA-LEFANSKRLALEKLNLDLESKLTKEKAEIOLAKLEAKALKLEOLAKOA 489  
 QY 374 KEL 376  
 DB 490 EEL 492

RESULT 9  
 SCA4\_RICAK STANDARD; PRT; 998 AA.  
 AC 09AIX9:  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ANTIGENIC HEAT-STABLE 120 KDA PROTEIN (Pst120) (120 KDA ANTIGEN)  
 GN SCA4 OR D.  
 OS Rickettsia akari.

CC Bacteria: Proteobacteria: alpha subdivision: Rickettsiales;  
 CC Rickettsiaceae; Rickettsiidae; Rickettsia.  
 CC MBIL\_TaxID=786;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sekeyova Z., Roux V., Raoult D.;  
 RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the  
 RT 'gene D' coding for an intracytoplasmic protein."  
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 CC -SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLY).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF213016; AAK30691.1; -  
 KW Antigen.  
 FT NON\_TER 1 1  
 FT NON\_TER 998 998  
 FT SEQUENCE 998 AA; 109328 MW; FCEB43AC62DE5BD5 CRC64;

Query Match 9.0%; Score 167.5; DB 1; Length 998;  
 Best Local Similarity 20.7%; Pred. No. 0.67;  
 Matches 102; Conservative 69; Mismatches 140; Indels 181; Gaps 22;

QY 5 NSG-----KDGWTMSANSADSV-----KGPNTLEINKKTTSDNAVL--LAVREVEL 49  
 DB 424 NAGLTKEKDGNTQIDLTNEAATAILNNEKEKQANFTLTKNVYNNALLPDRIKVARNAV 483  
 QY 50 LSSIDEIAAAGIKKIHQNGGLDPTENNHNGLAGAVAI STLKOKLDLGNKGL--KEK 107  
 DB 484 LETI-----KNQDTPDIEKSKMLAEVAIT-----LASENLTPKOK 520  
 QY 108 IDAKKCEFTTNKLEKHTDGLKEGVTDADAKKEALIKAN-STTKGAE--ELGKLFESV 164  
 DB 521 QOMLEKAVDVDSFKDQTSRAVAIDITG--AVIKSNLTOKKGMILAVGDQKVAS 575  
 QY 165 EVLSKAKEMLANSV-----KELTSP-----VAESKPKPSVNNSGKDGNT 206  
 DB 576 E-LSNAKEQQLGSLVAKKVEYKIIISPEQOOLMOQNLDKITAEDTKNDNTIEVOGILANP 634  
 QY 207 SANSADSV-----VKGPNLTLEISKRTES----- 230  
 DB 635 AENITAKTAIOGVTTKVLDSPTIAEIKGETLESTTKIVASPLNAVQDKDIYKMGGEAI 694  
 QY 231 -----NAVLAVKEVET-----LITS-----IDELAKAI 254  
 DB 695 ASHRTMAPTKIAALESEYETGAANSITDLEDKKMLTGLVDGIYEDKANPEITSEMKKAV 754  
 QY 255 GKIKNDVSLDNEADHNSLISGAVYLSNLTIKKISAIKSGELAKAEIKAKKCESEFTA 314  
 DB 755 SKGVDN-----STAPEDKOLAKPAAS-EAALDRA--TQNFTE 789  
 QY 315 KLGKHTDLGKEGYTDNNAKA-----IKTNNDKTGADE--LEKLFESV 359  
 DB 790 GLKQNLDEPKP--RDDIYNKAODIAVALKNVYTVLDANPEKREVSSEBVMNKTSILN 847  
 QY 360 NLSRAKEMLTN 371  
 DB 848 DISKIAIEKYN 859

RESULT 10  
 RBP2\_PLAVB STANDARD; PRT; 1251 AA.  
 ID RBP2\_PLAVB  
 AC 000799;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)



DR 01-OCT-1996 (Rel. 34, Last annotation update)  
DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).  
GN RBP2.  
OS Plasmodium vivax (strain Belen).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxId=31273;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92315338; PubMed=1617731;  
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
RT "A reticulocyte-binding protein complex of Plasmodium vivax  
mezozoites."  
RL Cell 69:1213-1226(1992).  
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
CC HUMAN RETICULOCYTE CELLS.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: M88098; AAA29744.1; -  
KW Malaria: Receptor; Membrane.  
FT NON\_TER 1  
FT NON\_TER 1251  
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;  
  
Query Match 9.0%; Score 166.5; DB 1; Length 1251;  
Best Local Similarity 22.6%; Pred. No. 0.95;  
Matches 91; Conservative 55; Mismatches 142; Indels 115; Gaps 17;  
  
QY 22 KGNPLTEIKKTTDSMAVLAVKEVALLSSIDEIAKAIGKKIHNNGLDPENNINGSI 81  
DB 774 KNTNELDVKKNIQDAKVALET-----LAHSDI-----DTKQKQSSKL 812  
QY 82 LA---GAVYISTLIKQKLDGLKNEGLKEKIDAKKCEFTFNKL---KEKHDTLGREGYT 135  
DB 813 IEMGNQIYIAKVLINQ-----YKNKISSIKSEAVSVKIGVSKHSHLSKTIIS 863  
QY 136 DADAKELT-----LKANGTKGAELGLKFESEVLSKAAKEMLVANSKELTSP 185  
DB 864 DKSVDNIILEKQTELQNLNRSFTQEKFTNTNSDSKLEI-----KTFESIKNALKTLEGE 919  
QY 186 VVAESPKKRSMVNNSGKQDNTSANSADSEVKGPNLTLEISKITTESAAVVLAVKEVETLLT 245  
DB 920 VNA-----LKASSDNEHVQSKSEPV-NPAUSEIKKEET-----DIDSLMT 959  
QY 246 SIDELAK-----AIGKKINDVSLDNEADHNSLISGAVL----- 280  
DB 960 ALDELKAKRTCEVSKYKLIKDTYVKEISDDELINTIEKNK---ATLAYIKKNYEDT 1015  
QY 281 ISNLT--KTSIKDSEGLKAEIEKAKKCESEFTAKLGEHTDLCKEGVTDNNAKAIL 1038  
DB 1016 VQDVLTLLNHFMTKQVSNHEPTNFDKSNKSEBELTKAVVDSKTIISK-----LKGVII 1068  
QY 339 KTN-----NDKTKGDELKFESEVKNLSKAAKEM--LNNSVK 374  
DB 1069 EYNEENTEMNTIESSAKETALYNELNKKTKTSLENIYOTSNEYA 1111  
  
RESULT 11  
MYS\_AEOIR STANDARD; PRT: 1938 AA.  
AC P24733;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MYOSIN HEAVY CHAIN, STRIATED MUSCLE.

OS Aequipecten irradians (Bay scallop).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;  
OC Pectinoidea; Pectinidae; Argopecten.  
OX NCBI\_TaxId=31199;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92011595; PubMed=1917970;  
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;  
RT "Complete primary structure of a scallop striated muscle myosin heavy  
RT chain. Sequence comparison with other heavy chains reveals regions  
RT that might be critical for regulation."  
RL J. Biol. Chem. 266:18469-18476(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Adductor muscle;  
RX MEDLINE=9108319; PubMed=2263488;  
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;  
RT "Nucleotide sequence of full length cDNA for a scallop striated  
RT muscle myosin heavy chain."  
RL Nucleic Acids Res. 18:7158-7158(1990).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.  
RX MEDLINE=94173332; PubMed=8127365;  
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,  
RT Szent-Gyorgyi A.G., Cohen C.;  
RT "Structure of the regulatory domain of scallop myosin at 2.8-A  
RT resolution."  
RL Nature 368:306-312(1994).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.  
RX MEDLINE=96419133; PubMed=8805510;  
RA Houdusse A., Cohen C.;  
RT "Structure of the regulatory domain of scallop myosin at 2-A  
RT resolution: implications for regulation."  
RL Structure 4:21-32(1996).  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: X55714; CAA39247.1; -  
DR PIR: S1357; S13587.  
DR PIR: A40997; A40997.  
DR PDB: 1SCM; 30-APR-94.  
DR PDB: 1WDC; 11-JUL-96.  
DR InterPro: IPR000048; IO.  
DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR002017; Spectrin.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ; 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF01576; myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR PRODOM: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ; 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ; 1.  
KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.

FT DOMAIN 1 777 MYOSIN HEAD-LIKE.  
 FT DOMAIN 778 805 IQ.  
 FT DOMAIN 836 1938 ROD-LIKE TAIL (S2 AND LAM DOMAINS).  
 FT DOMAIN 836 1938 COILED COIL (POTENTIAL).  
 FT NP\_BIND 176 183 ATP (BY SIMILARITY).  
 FT MOD\_RES 693 693 ALKYLATION (SH-1) (BY SIMILARITY).  
 FT MOD\_RES 703 703 ALKYLATION (SH-2) (BY SIMILARITY).  
 FT HELIX 778 821  
 FT TURN 822 823  
 FT HELIX 825 833  
 FT TURN 834 835  
 SO SEQUENCE 1938 AA; 222821 MW; ASCCE4127D1A4896 CRC64;

Query Match 8.9%; Score 165.5; DB 1; Length 1938;  
 Best Local Similarity 22.3%; Pred. No. 1.7;  
 Matches 88; Conservative 59; Mismatches 167; Indels 81; Gaps 12;

18 DESVKGPNLEINKKTTSDNAVLAAVEEALLSIDELAIAKIGKIHONGLDT----73  
 DB DEEDAAADLGIRKKMEADNA-----NLKKDIGLENTLOKAEODKAKHNDQISTLQGE 978  
 QY 74 ---ENNHGSLAGAVYISTLIKOKLDGLKNEGLKKEIDAAKCSFTTNKLEKHTDGL 130  
 DB 979 ISQDDEHIGLNKEKALEANAKTSDSLOAE--EDKCNHLNKLAKLEQALDELNDLE 1036  
 QY 131 KEVYTDAAKEAILKANGTYTKGAELGKLFESVEVLSKAKEMLANVKEITSPVAES 190  
 DB 1037 REKRVKQDVEKA-----KRVKQDLSKQEVNEDLEHVKRELEEN-----VRKK 1080  
 QY 191 PKRPSWVNGSKGNTSANSADSEVK--GNPLEISKRT-TSNNAVLAKEVEETLLSTI 247  
 DB 1081 EATISSLSNLEDEQNLVLSOLKRIEQLARIELEELAEARNRAKVEKORAEINREL 1140  
 QY 248 DELAKAIGK-----KIRNDV---SLDNEADHNGSLISGAVLISNL 284  
 DB 1141 EELGERLDEAGATSAQTELNKREAEILKIRDLSEASLQHEA-----ISGL 1189  
 QY 285 ITRKISAIIDSGELKEIKAKKCSFEETAKLKGHTDGLKGVYDNDKAKAILKTNNDR 344  
 DB 1190 RKKHODANEMADVOVDOLVKSKLEKDKDLKREMDLESD-----MTHMK 1237  
 QY 345 TKGADE-LKLFESVYNLSKAKEMITNSVKELTS 378  
 DB 1238 NKGCEKVKKQFESQMSDLNMLEDSQRSINELQS 1272  
 QY 12  
 USOL\_YEAST STANDARD: PRT; 1790 AA.  
 AC P25386;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DE 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.  
 GN USOL OR INT1 OR YDL058W.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=91185402; PubMed=2010462;  
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
 RA Yamaseki M.;  
 RT "A cytoskeleton-related gene, usol, is required for intracellular  
 RT protein transport in Saccharomyces cerevisiae.";  
 RL J. Cell Biol. 113:245-260(1991).  
 RN [2]  
 RP SEQUENCE OF 782-1790 FROM N.A.  
 RA Hostetter M.K., Herman D.J., Bendel C.M., McCellan M., Tao N.,  
 RA Kendrick K.E.;

RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -! FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC -! SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -! DOMAIN: THE ROD-LIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE. COMPOSED  
 CC OF AN HEPAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC -! SIMILARITY: BELONGS TO THE YDP/USOL/YBL047C FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X54378; CA38253.1; -;  
 DR EMBL; L03188; AAB00143.1; -;  
 DR EMBL; U53668; AAB6659.1; -;  
 DR PIR; A38455; A38455.  
 DR HSSP; P80220; IDIP.  
 DR SGD; S0002216; USOL.  
 DR InterPro; IPR002017; Spectrin.  
 DR Transprot; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
 KW DOMAIN 1 724  
 FT DOMAIN 725 1790  
 FT DOMAIN 465 487 COILED COIL (POTENTIAL).  
 FT DOMAIN 991 1790 CHARGED (HYPER-HYDROPHILIC).  
 FT DOMAIN 1172 1786 DISPENSABLE FOR THE PROTEIN FUNCTION.  
 FT CONFLICT 847 847 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 924 924 G -> E (IN REF. 2).  
 FT CONFLICT 1253 1253 E -> K (IN REF. 2).  
 FT CONFLICT 1319 1319 V -> I (IN REF. 2).  
 FT CONFLICT 1461 1461 I -> V (IN REF. 2).  
 FT CONFLICT 1581 1581 N -> S (IN REF. 2).  
 FT CONFLICT 1600 1600 G -> S (IN REF. 2).  
 FT CONFLICT 1661 1661 I -> V (IN REF. 2).  
 FT CONFLICT 1772 1772 R -> S (IN REF. 2).  
 SO SEQUENCE 1790 AA; 206424 MW; 6CE2B210E9FD4818 CRC64;

Query Match 8.8%; Score 164; DB 1; Length 1790;  
 Best Local Similarity 21.6%; Pred. No. 1.8;  
 Matches 101; Conservative 75; Mismatches 186; Indels 106; Gaps 17;

QY 3 CNNSGKDG-----NTSANSADSEYKGNPLEINKKTTSDNAVLAAVEEALLSS 52  
 DB 929 CNLSKEKEHISELYEYKSRFQSHDNLV--AKTEKULSLANNYKMQA--ENESLIKA 984  
 QY 53 IDEIAAKAGKRIHONGLD--TENNHNGSLAGAV-----AISTL-----IKQ 95  
 DB 985 VESKNESSIQLSNQNKDMSQSKENFQIRGSIENINQIKTTISDLQOTKEELISK 1044  
 QY 96 LDKLNKG-----LKEIDAAKCSFTTNKLE-----KHYDLKREGVTD 136  
 DB 1045 SSSSDVEESQISLLEKLETAATTANDENVKVISLFTYRRELEAEALAAVKNLNELET 1104  
 QY 137 ADAKPAIILK-----ANGTKTK-----GAEILGKLFESVAVLSKAA 171  
 DB 1105 LETSKALKKEVENEHLEKEIKQLEKATETKQOOLNSRLANLESLKEHEHDLAAQKTY 1164  
 QY 172 KEMLANVKEITSPV--VAESPKRPSWVNGSKGNTSANSADSEYKGNPLEISKRT 229  
 DB 1165 EQLANKEROYNEISQNLDEITSTQOENESIKKKNDLEGEVAKAMKSTBEQSINLKSE 1224  
 QY 230 SNAVYLAKEY-----ETLLTSDILAIAIGKIKNDVSLDNEADHNGSLISGAVLISNLI 285

```

DB 1225 IALNQLIKELKKNKTNSLESIKSVSESEVYKIKELODECFEKESE-----L 1277
QY 286 TKKISAIKSG-----ELKAEIEKAKCSEBFAKLK---GEHTDLCEGVTDNNAKAIL 338
DB 1278 EKLKASEKSNKYVELKSEKIKELDAKTTELKIOLEKTLNLSKAKESSELSRLK 1337
QY 339 KKNNDKTKADE-LEKL-----FESVKLSKAKEMLNVSKE 375
DB 1338 KTSSEKRNKAEDEKLEKLNKIOIKNAEFERKRLNLSGSSITTOEYSE 1385

RESULT 13
MYS2_DICDI
ID MYS2_DICDI STANDARD; PRT: 2116 AA.
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
CN MHCA.
CM Dicyostellium discoidium (Slime mold).
OX Eukaryota; Mycetozoa; Dictyostelida; Dictyostellium.
RN NCBI_TaxID=44689;
RN [1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warlick H.M., de Lozanne A., Leitwand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
RN Dicyostellium discoidium."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RN RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Velometer D., Schleicher M., Grabatin B., Wippler J.,
RN Gerlach G.;
RT "Replacement of threonine residues by serine and alanine in a
RN phosphorylatable heavy chain fragment of Dicyostellium myosin II."
RL FEBS Lett. 269:239-243(1990).
RN [3]
RN RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Magle G., Noegel A., Scheel J., Gerlach G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RN Dicyostellium myosin heavy chain."
RL FEBS Lett. 227:71-75(1988).
RN [4]
RN RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RN Rayment I.;
RT "X-ray structures of the myosin motor domain of Dicyostellium
RN discoidium complexed with MgADP.Befx and MgADP.ALFA."
RL Biochemistry 34:8960-8972(1995).
RN [5]
RN RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RN truncated head of Dictyostellium discoidium myosin to 2.7-A
RN resolution."
RL Biochemistry 34:8973-8981(1995).
RN [6]
RN RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RN Dicyostellium discoidium myosin motor domain to 1.9-A resolution."
RL Biochemistry 35:5404-5417(1996).
RN [7]
RN RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX MEDLINE=97452580; PubMed=9305951;

```

```

RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammaS, and MgADPPNP complexes
RN of the Dicyostellium discoidium myosin motor domain."
RL Biochemistry 36:11619-11628(1997).
RN [8]
RN RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RN of Mg.2(3')-O-(N-methylanthranilloyl) nucleotides bound to the
RN Dicyostellium discoidium myosin motor domain."
RL J. Mol. Biol. 274:394-407(1997).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC CORTEX.
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN BE FURTHER
CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHARED
CC SUBFRAGMENT (S2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: DICTYOSTELLUM MYOSIN II HAS NO K(2)DTA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14628; AAA33227.1; -
DR PIR: A26655; A26655.
DR PIR: S00250; S00250.
DR PDB: 1MMA: 03-DEC-97.
DR PDB: 1MMD: 17-AUG-96.
DR PDB: 1MMG: 03-DEC-97.
DR PDB: 1MMN: 03-DEC-97.
DR PDB: 1MND: 17-AUG-96.
DR PDB: 1MNE: 17-AUG-96.
DR PDB: 1VOM: 23-DEC-96.
DR PDB: 1LVK: 28-JAN-98.
DR Dictydb: DDO1008; mhca.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ. 1.
DR Pfam: PF00663; myosin_head; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ. 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
DR Myosin: Coiled coil; Actin-binding; 3D-structure;
KW Calmodulin-binding; Methylation; ACTIN-BINDING; PHOSPHORYLATION.
FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
FT DOMAIN 762 791 IQ.
FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
FT NF_BIND 179 186 ATP.
FT DOMAIN 638 660 ACTIN-BINDING.
FT DOMAIN 738 752 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).

```



```

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC
CC -1- PRIM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEMOROSIN (LMN) AND 1 HEAVY MEMOROSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHADED
CC SUBFRAGMENT (S2).
CC
CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC ELEFANS.
CC
CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC
CC EMBL: J01050; AAA28124.1; -.
CC DR EMBL: V01494; CAA24738.1; -.
CC DR PIR: A02992; MMKW.
CC
CC HSSP: P08799; 1MND.
CC
CC InterPro: IPR002928; Myosin_tail.
CC DR InterPro: IPR001609; myosin_head.
CC DR Pfam: PF00063; myosin_head; 1.
CC DR Pfam: PF01576; Myosin_tail; 1.
CC DR PRINTS: PR00193; MYOSINHEAVY.
CC DR PRODOM: PD000355; myosin_head; 1.
CC
CC SMART: SM00242; MYSC: 1.
CC
CC Myosin: Muscle protein; coiled coil; thick filament; Actin-binding;
CC ATP-binding; Methylation; Alkylation; Multigene family.
CC
CC FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
CC FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
CC FT DOMAIN 1165 1176 ALPHA-HELICAL TAILPIECE (S2).
CC FT DOMAIN 1165 1966 HINGE.
CC FT DOMAIN 1165 1966 LIGHT MEMOROSIN (LMN).
CC FT NP_BIND 177 184 ATP (BY SIMILARITY).
CC FT DOMAIN 655 687 ACTIN-BINDING.
CC FT DOMAIN 769 783 ACTIN-BINDING.
CC FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
CC FT MOD_RES 705 705 ALKYLATION (SH-1).
CC FT MOD_RES 715 715 ALKYLATION (SH-2).
CC FT CONFLICT 1337 1337 E -> R (IN REF. 2).
CC FT CONFLICT 1880 1880 I -> L (IN REF. 2).
CC
CC SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CRC64;

```



Tues Mar 19 10:57:46 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:46 ; Search time 68.77 Seconds  
(without alignments)  
418.700 Million cell updates/sec

Title: US-09-596-746a-28

Perfect score: 1859

Sequence: 1 MACNNNSGRKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	50.8	210	2	G70218
2	938	50.5	210	2	S69927
3	879	47.3	210	2	I40144
4	853	45.9	177	2	I40129
5	774.5	41.7	211	2	I40145
6	773.5	41.6	193	2	S70279
7	700.5	37.7	193	2	S70287
8	695.5	37.4	211	2	S69918
9	692.5	37.3	211	2	I40277
10	690.5	37.1	193	2	S70280
11	686.5	36.9	211	2	I40268
12	686.5	36.9	190	2	I40273
13	685.5	36.9	209	2	I40273
14	683.5	36.8	191	2	S70278
15	683.5	36.8	191	2	I40153
16	675	36.3	194	2	S70277
17	674.5	36.3	209	2	S69917
18	670	36.0	212	2	S70254
19	666.5	35.9	193	2	S70276
20	664.5	35.7	193	2	S70274
21	664.5	35.7	211	2	I40278
22	664	35.7	212	2	S69922
23	662.5	35.6	191	2	S70284
24	662.5	35.6	193	2	S70265
25	662	35.6	194	2	S70268
26	661.5	35.6	211	2	S69932
27	657.5	35.4	209	2	I40142
28	655	35.2	212	2	I40279
29	650	35.0	212	2	I40143

30	646	34.7	214	2	S69916	outer surface prot
31	641.5	34.5	193	2	S70286	outer surface prot
32	638	34.3	194	2	S70289	outer surface prot
33	635.5	34.2	191	2	S70288	outer surface prot
34	635.5	34.2	203	2	I40108	outer surface prot
35	631	33.9	212	2	S20543	outer surface prot
36	630.5	33.9	209	2	I40281	outer surface prot
37	630.5	33.9	209	2	I40285	outer surface prot
38	630	33.9	192	2	S70285	outer surface prot
39	628	33.8	210	2	S69925	outer surface prot
40	627	33.7	210	2	S69923	outer surface prot
41	627	33.7	212	2	S69921	outer surface prot
42	622.5	33.5	209	2	S69926	outer surface prot
43	622	33.5	210	2	S69920	outer surface prot
44	619	33.3	210	2	I40272	outer surface prot
45	617.5	33.2	193	2	S70259	outer surface prot

## ALIGNMENTS

RESULT 1  
G70218  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000  
C:Accession: G70218; I40269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kevlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; M0ID:98065943  
A:Accession: G70218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <KLE>  
A:Cross-references: GB:AE000792; NID:93253098; PIDN:AAC66329.1; PID:92689901; TIGR:BB  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu  
stricto.  
A:Reference number: I40269; M0ID:96025162  
A:Accession: I40269  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49497; NID:9707092; PIDN:BA08457.1; PID:9769684  
R:Daubits-Heipke, S.; Fuchs, R.; Mottz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os  
pC) of Borrelia burgdorferi.  
A:Reference number: S37726; M0ID:93268136  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69596; NID:9311391; PIDN:CAA49306.1; PID:9311392  
R:Li, Y.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
sease spirochetes.  
A:Reference number: S70253; M0ID:96296448  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:I42887; NID:9858715; PIDN:AAB36959.1; PID:91695212  
A:Experimental source: strain Ip2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.8%; Score 945; DB 2; Length 210;  
 Best Local Similarity 97.9%; Pred. No. 4e-37;  
 Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNPTLEINKITDSNAVLAVKVEVALLSSIDEIAKA 60  
 DB 17 ISCNNSGKDGNTSANSADSVKGNPTLEISKITDSNAVLAVKVEVALLSSIDEIAKA 60  
 QY 61 IGRKHQNNGLDTEENNHNHNSLAGAYASTLKOKLDLKNKGKLEKIDAKKCEFTFN 76  
 DB 77 IGRKHQNNGLDTEENNHNHNSLAGAYASTLKOKLDLKNKGKLEKIDAKKCEFTFN 120  
 QY 121 KLEKEHTDGLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 180  
 DB 137 KLEKEHTDGLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 196  
 QY 181 ELTSPVAESPKRP 194  
 DB 197 ELTSPVAESPKRP 210

RESULT 2  
 S69927  
 outer surface protein C precursor - Lyme disease spirochete (strain PKA)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 A:Variety: strain PKA  
 C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
 C:Accession: S69927; S72669  
 R:Jauris-Helpe, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.  
 J. Clin. Microbiol. 33: 1860-1866, 1995  
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia  
 A:Reference number: I40047; MUID:95395018  
 A:Accession: S69927  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-210 <JAU>  
 A:Cross-references: EMBL:X69589  
 R:Jauris, S.  
 A:Experimental source: strain PKA  
 A:Submitted to the EMBL Data Library, February 1994  
 A:Reference number: S72669  
 A:Accession: S72669  
 A:Molecule type: DNA  
 A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAW>  
 A:Cross-references: EMBL:X69589  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.5%; Score 938; DB 2; Length 210;  
 Best Local Similarity 96.9%; Pred. No. 8.5e-37;  
 Matches 188; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNPTLEINKITDSNAVLAVKVEVALLSSIDEIAKA 60  
 DB 17 ISCNNSGKDGNTSANSADSVKGNPTLEISKITDSNAVLAVKVEVALLSSIDEIAKA 60  
 QY 61 IGRKHQNNGLDTEENNHNHNSLAGAYASTLKOKLDLKNKGKLEKIDAKKCEFTFN 76  
 DB 77 IGRKHQNNGLDTEENNHNHNSLAGAYASTLKOKLDLKNKGKLEKIDAKKCEFTFN 120  
 QY 121 KLEKEHTDGLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 180  
 DB 137 KLEKEHTDGLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 196  
 QY 181 ELTSPVAESPKRP 194  
 DB 197 ELTSPVAESPKRP 210

RESULT 3  
 S140144

outer surface protein C precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 C:Accession: I40144; S70282  
 R:Stevenson, B.; Barthold, S.W.  
 FEMS Microbiol. Lett. 124, 367-372, 1994  
 A:Title: Expression and sequence of outer surface protein C among North American isol  
 A:Reference number: I40143; MUID:95154673  
 A:Accession: I40144  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-210 <RES>  
 A:Cross-references: EMBL:U04281; NID:9434663; PIDN:AA03297.1; PID:9434664  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70282  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 19-210 <LIV>  
 A:Cross-references: EMBL:L42893; NID:958721; PIDN:AA037001.1; PID:91695218  
 C:Experimental source: strain 297  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 47.3%; Score 879; DB 2; Length 210;  
 Best Local Similarity 99.4%; Pred. No. 4.3e-34;  
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NNSGKDGNTSANSADSVKGNPTLEISKITDSNAVLAVKVEVALLSSIDEIAKAIGK 257  
 DB 20 NNSGKDGNTSANSADSVKGNPTLEISKITDSNAVLAVKVEVALLSSIDEIAKAIGK 79  
 QY 258 IKNVSLDNEADHNGSLISGAYLISLITKISAIKDSGELKAELEKAKKCEFTATLK 317  
 DB 80 IKNVSLDNEADHNGSLISGAYLISLITKISAIKDSGELKAELEKAKKCEFTATLK 139  
 QY 318 GEHTDLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 377  
 DB 140 GEHTDLGEGVTDADAKKAEILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSYK 199  
 QY 378 S 378  
 DB 200 S 200

RESULT 4  
 I40129  
 outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
 C:Accession: I40129; S54199  
 R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
 J. Bacteriol. 177, 3036-3044, 1995  
 A:Title: Evolution of the Borrelia burgdorferi outer surface protein ospC.  
 A:Reference number: I40104; MUID:95286481  
 A:Accession: I40129  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-177 <RES>  
 A:Cross-references: EMBL:X84783; NID:9793825; PIDN:CAA59254.1; PID:9793826  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 45.9%; Score 853; DB 2; Length 177;  
 Best Local Similarity 98.3%; Pred. No. 5.5e-33;  
 Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;



OY 11 NTSANSADESVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAKIGKHQNNNG 70  
 DB 1 NTSANSADESVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAKIGKHQNNNG 60  
 OY 71 LDTENNHNHSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNKLEKHTDGL 130  
 DB 61 LDTENNHNHSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNKLEKHTDGL 120  
 OY 131 KCGVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVKELTSPV 187  
 DB 121 KCGVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVKELTSPV 177  
 RESULT 5  
 I40145  
 outer surface protein C precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 C:Accession: I40145  
 R:Stevenson, B.; Barthold, S.W.  
 PMS Microbiol. Lett. 124, 367-372, 1994  
 A:Title: Expression and sequence of outer surface protein C among North American isolate  
 A:Reference number: I40143; MUID:95154673  
 A:Accession: I40145  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-211 <RES>  
 A:Cross-References: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.7%; Score 774.5; DB 2; Length 211;  
 Best Local Similarity 81.0%; Pred. No. 2.7e-29;  
 Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

OY 1 MACNNSGKDCN-TSANSADESVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 59  
 DB 17 ISCNNSGKDCNASTNPADSVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 76  
 OY 60 AIGKRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFT 119  
 DB 77 AIGKRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFT 136  
 OY 120 NKLKHTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 179  
 DB 137 NKLKHTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 196  
 180 KELTSPPVAESPCKP 194  
 197 KELTSPPVAESPCKP 211

RESULT 6  
 S70279  
 outer surface protein C - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70279  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70279  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIV>  
 A:Cross-References: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223  
 A:Experimental source: strain 25015  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.6%; Score 773.5; DB 2; Length 193;  
 Best Local Similarity 81.9%; Pred. No. 2.7e-29;  
 Matches 158; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

OY 3 CNNSGKDCN-TSANSADESVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 61  
 DB 1 CNNSGKDCNASTNPADSVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 60  
 OY 62 GKRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNK 121  
 DB 61 GKRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNK 120  
 OY 122 LKERTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 181  
 DB 121 LKERTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 180  
 OY 182 LTPSPVAESPCKP 194  
 DB 181 LTPSPVAESPCKP 193

RESULT 7  
 S70287  
 outer surface protein C - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70287  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70287  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIV>  
 A:Cross-References: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220  
 A:Experimental source: strain 28354  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.7%; Score 700.5; DB 2; Length 193;  
 Best Local Similarity 75.1%; Pred. No. 5.9e-26;  
 Matches 145; Conservative 17; Mismatches 30; Indels 1; Gaps 1;

OY 3 CNNSGKDCN-TSANSADESVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 62  
 DB 1 CNNSGKDCNASTNPADSVKGNPLTEINKKTTSDSNVAVLAVKEVEALLSIDETIAAK 60  
 OY 63 KRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNK 121  
 DB 61 KRTKHONNGLDTENNHNSLAGAVAIATSTLIKOKLDGLKNGELKEKIDAAKCSSEFTTNK 120  
 OY 122 LKERTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 181  
 DB 121 LKERTDGLKGEVTDADAKKAILKANGTKTKGAEBELKLFESVEVLSKAEMKANSVK 180  
 OY 182 LTPSPVAESPCKP 194  
 DB 181 LTPSPVAESPCKP 193

RESULT 8  
 S69918  
 outer surface protein C precursor - Lyme disease spirochete (strain p8e)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 A:Variety: strain p8e  
 C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
 C:Accession: S69918; S72674; I40103  
 R:tautis-Helpke, S.; Diegl, G.; Prenc-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek

J. Clin. Microbiol. 33, 1860-1866, 1995  
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia burgdorferi*  
 A:Reference number: 140047, MID:9595018  
 A:Accession: S69918  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-211 <ORF>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDD:CAA57242.1; PID:9872022  
 A:Experimental source: strain pRe  
 R:Roessler, D.  
 Submitted to the EMBL Data Library, September 1994  
 A:Reference number: S72674  
 A:Accession: S72674  
 A:Molecule type: DNA  
 A:Residues: 1-152, 'E', 154-211 <ORF>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDD:CAA57242.1; PID:9872022  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.4%; Score 695.5; DB 2; Length 211;  
 Best Local Similarity 77.8%; Pred. No. 1.1e-25;  
 Matches 152; Conservative 12; Mismatches 29; Indels 3; Gaps 3;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKA 60  
 DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKA 60  
 QY 61 IGRKHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-NEGLEKIDAKKCEFTT 119  
 DB 76 IGRKHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-NEGLEKIDAKKCEFTT 119  
 QY 120 NKLEKHTDLGKEGYTDADAEALIKANGT-KTKGAELGLFESEVLSKAKEMLANS 178  
 DB 136 TKLKDHAQLGIGVYTDNNAKRAILKANAGKDKGVELEKLSGLSLSKAKEMLANS 195  
 QY 179 VKELTSPVAESPCKP 194  
 DB 196 VKELTSPVAESPCKP 211

RESULT 9  
 140277  
 outer surface protein C precursor - *Borrelia afzelii*  
 C:Species: *Borrelia afzelii*  
 C:Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
 C:Accession: 140277  
 R:Fukunaga, M.; Hamase, A.  
 J. Clin. Microbiol. 33, 2415-2420, 1995  
 A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato  
 A:Reference number: 140269; MID:96025162  
 A:Accession: 140277  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-211 <RES>  
 A:Cross-references: GB:DA9501; NID:9707096; PIDD:BA08461.1; PID:9769688  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.3%; Score 692.5; DB 2; Length 211;  
 Best Local Similarity 74.5%; Pred. No. 1.1e-25;  
 Matches 146; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKA 59  
 DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKA 59  
 QY 60 AIGKHHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-EGLEKIDAKKCEFTT 118  
 DB 77 AIGKHHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-EGLEKIDAKKCEFTT 118  
 QY 119 TNKLEKHTDLGEGYTDADAEALIKANGT-KTKGAELGLFESEVLSKAKEMLANS 178

DB 136 TKLKDSDNADLCKHATADDSKEALIKNTGKTGAKELBELPKSVESLSKAKEMLANS 195  
 QY 179 VKELTSPVAESPCKP 194  
 DB 196 VKELTSPVAESPCKP 211

RESULT 10  
 S70280  
 outer surface protein C - Lyme disease spirochete  
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70280  
 R:Livey, J.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in *OspC* variation in Lyme disease  
 A:Reference number: S70285; MID:96296448  
 A:Accession: S70280  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIV>  
 A:Cross-references: EMBL:L42868; NID:9858735; PIDD:AA837011.1; PID:91695226  
 A:Experimental source: strain ZS7  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.1%; Score 690.5; DB 2; Length 193;  
 Best Local Similarity 77.8%; Pred. No. 1.7e-25;  
 Matches 151; Conservative 10; Mismatches 30; Indels 3; Gaps 3;

QY 3 CNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKAIG 62  
 DB 1 CNSGKDGNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSSIDEIAKAIG 59  
 QY 63 KKHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-NEGLEKIDAKKCEFTT 121  
 DB 60 KKHONNGLDTENNHNSLAGAVALSTLKOKLDGLK-NEGLEKIDAKKCEFTT 121  
 QY 122 LKEKHTDLGKEGYTDADAEALIKANGT-KTKGAELGLFESEVLSKAKEMLANSV 180  
 DB 120 LKDNHQAQLGIGVYTDNNAKRAILKANAGKDKGVELEKLSGLSLSKAKEMLANSV 179  
 QY 181 ELTSPVAESPCKP 194  
 DB 180 ELTSPVAESPCKP 193

RESULT 11  
 140268  
 outer surface protein C precursor - Lyme disease spirochete  
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 26-May-2000  
 C:Accession: 140268  
 R:Margolis, N.; Hogan, D.; Cieplak, W.  
 Gene 143, 105-110, 1994  
 A:Title: Homology between *Borrelia burgdorferi* *OspC* and members of the family of *Borrelia burgdorferi*  
 A:Reference number: 140268; MID:94259285  
 A:Accession: 140268  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-211 <RES>  
 A:Cross-references: GB:L25413; NID:9495735; PIDD:AAA22956.1; PID:9495736  
 C:Genetics:  
 A:Gene: ospC  
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 686.5; DB 2; Length 211;  
 Best Local Similarity 73.8%; Pred. No. 2.9e-25;  
 Matches 144; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 60  
 DB 17 ISCNNSGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 76  
 OY 61 ICKKIHNNGLDTENNHNSGLAGAYASTLIKOKLDGLKNEG-LKEKIDAAKCSSEFTT 119  
 DB 77 ICKKIHNNGLDTENNHNSGLAGAYASTLIKOKLDGLKNEG-LKEKIDAAKCSSEFTT 136  
 OY 120 NKLEKHTDLGREGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 179  
 DB 137 KRLDSNADLGKHNATDADSKAAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 196  
 OY 180 KELTSPVVAESPCKP 194  
 DB 197 KELTSPVVAESPCKP 211

RESULT 12  
 S70273  
 outer surface protein C - Lyme disease spirochete  
 Species: Borrelia burgdorferi (Lyme disease spirochete)  
 Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70273  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70273  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-190 <LIV>  
 A:Cross-References: EMBL:L42870; NID:9858737; PIDN:AAB37013.1; PID:g1695228  
 A:Experimental source: strain VSDA  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 686; DB 2; Length 190;  
 Best Local Similarity 75.6%; Pred. No. 2.7e-25;  
 Matches 146; Conservative 19; Mismatches 24; Indels 4; Gaps 3;  
 OY 3 CANSNGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 62  
 DB 1 CANSNG--GDTASTNPDESAGPDLTVISKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 57  
 OY 63 KTIHONGGLDTENNHNHNSGLAGAYASTLIKOKLDGLK-NEGLEKIDAAKCSSEFTT 121  
 DB 58 KTIHONGGLDTENNHNHNSGLAGAYASTLIKOKLDGLKSEGKAEIAEAKKCSSEFTT 117  
 OY 122 LKEKHTDLGREGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 181  
 DB 118 LKEKHTDLGREGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 177  
 OY 182 LTPSPVVAESPCKP 194  
 DB 178 LTPSPVVAESPCKP 190  
 RESULT 13  
 I40273  
 outer surface protein C precursor - Borrelia afzelii  
 C:Species: Borrelia afzelii  
 C:Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
 C:Accession: I40273  
 R:Fukunaga, M.; Hamase, A.  
 J. Clin. Microbiol. 33, 2415-2420, 1995  
 A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la  
 A:Reference number: I40269; MUID:96025162  
 A:Accession: I40273  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Residues: 1-209 <RES>  
 A:Cross-References: GB:D49379; NID:q1041107; PIDN:BA08377.1; PID:g1041108  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 685.5; DB 2; Length 209;  
 Best Local Similarity 74.4%; Pred. No. 3.2e-25;  
 Matches 145; Conservative 21; Mismatches 26; Indels 3; Gaps 3;  
 OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 60  
 DB 17 ISCNNSGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 75  
 OY 61 ICKKIHNNGLDTENNHNSGLAGAYASTLIKOKLDGLKNEG-LKEKIDAAKCSSEFTT 119  
 DB 76 ICKKIHNNGLDTENNHNSGLAGAYASTLIKOKLDGLKNEG-LKEKIDAAKCSSEFTT 134  
 OY 120 NKLEKHTDLGREGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 179  
 DB 135 KRLDSNADLGKHNATDADSKAAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 194  
 OY 180 KELTSPVVAESPCKP 194  
 DB 195 KELTSPVVAESPCKP 209

RESULT 14  
 S70278  
 outer surface protein C - Lyme disease spirochete  
 Species: Borrelia burgdorferi (Lyme disease spirochete)  
 Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70278  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70278  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-191 <LIV>  
 A:Cross-References: EMBL:L42871; NID:9858738; PIDN:AAB37014.1; PID:g1695229  
 A:Experimental source: strain VSD461  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.8%; Score 683.5; DB 2; Length 191;  
 Best Local Similarity 75.1%; Pred. No. 3.5e-25;  
 Matches 145; Conservative 19; Mismatches 26; Indels 3; Gaps 3;  
 OY 3 CANSNGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 62  
 DB 1 CANSNGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 59  
 OY 63 KTIHONGGLDTENNHNHNSGLAGAYASTLIKOKLDGLK-EGLEKIDAAKCSSEFTT 121  
 DB 60 KTIHONGGLDTENNHNHNSGLAGAYASTLIKOKLDGLKSEGKAEIAEAKKCSSEFTT 118  
 OY 122 LKEKHTDLGREGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 181  
 DB 119 LKDSNADLGKHNATDADSKAAILKANGTKTGAEELGKLFESVEVLSKAAREMLANSV 178  
 OY 182 LTPSPVVAESPCKP 194  
 DB 179 LTPSPVVAESPCKP 191  
 RESULT 15  
 I40153  
 outer surface protein C - Lyme disease spirochete (fragment)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 26-May-2000

C:Accession: I40153  
 R:Funf, B.P.; MCHugh, G.L.; Leong, J.M.; Steere, A.C.  
 Infect. Immun. 62, 3213-3221, 1994  
 A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in I  
 A:Reference number: I40153; MUID:94314437  
 A:Accession: I40153  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-191 <RES>  
 A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.8%; Score 683.5; DB 2; Length 191;  
 Best Local Similarity 74.9%; Pred. No. 3.5e-25;  
 Matches 143; Conservative 17; Mismatches 30; Indels 1; Gaps 1;  
 QY 5 NSGKGNFSANSADSVKGNLTETNKKITDSDNAVLAVKEVEALLSSIDEIAKAIGKK 64  
 Db 1 NSGKGNFSANSADSVKGNLTETNKKITDSDNAVLAVKEVEALLSSIDEIAKAIGKK 60  
 QY 65 IHQNGLDPTENNHNHNSILAGAYATSTLLKQKLDGLKN-EGIKETIDAKKCSFTTNKLR 123  
 Db 61 IQQNGGLAVEAGHNGTLLAGAYTTISKLITQKLDGLKNSEKLEKIEENAKKCSDEFTKLE 120  
 QY 124 EKHTDLGEGVTADAKAEAILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVKELT 183  
 Db 121 GEHQQLGIENVTDENAKKAILITDAAKDKGAELKLEKFAVENLAKRAKEMLANSVKELT 180  
 QY 184 SPVVAESPCKP 194  
 Db 181 SPVVAESPCKP 191

Search completed: March 18, 2002, 09:56:47  
 Job time: 370 sec



Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

## RESULT 2

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180  
QY 183 TSPVVAESPCKP 194  
181 TSPVVAESPCKP 192

```

DB 67 IGGKIHONNGLDTEENNHNHNSGLAGAYAISTLIKOKLDGKLNKKEKIDAKKCEFTFTN 126
QY 121 KLEKHTDLDKEGVTDADAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 127 KLEKHTDLDKEGVTDADAEALITKNGTKGAEELGKLFESVEVLSKAKEMLANSVK 186
QY 181 ELTSPV 187
DB 187 ELTSPV 193

RESULT 5
Q9ROR9 PRELIMINARY: PRT: 200 AA.
AC Q9ROR9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS PC.
DB Borrelia burgdorferi (Lyme disease spirochete).
Plasmid cp26.
Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RN
RP SEQUENCE FROM N.A.
RC STRAIN=267;
RX MEDLINE=20002545; Pubmed=10531219;
RA Hofmeister E.K., Glas G.E., Childs J.E., Persing D.H.;
RT "Population dynamics of a naturally occurring heterogeneous mixture of
  Borrelia burgdorferi clones."
  Infect. Immun. 67:5709-5716(1999).
RL EMBL: AF074464; AAD23911.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BB6B6B CRC64;

Query Match 48.0%; Score 893; DB 2; Length 200;
Best Local Similarity 97.8%; Pred. No. 4,4e-34;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTLEINKKITDSNAVLLAVKEVALLSIDEIATAKA 60
DB 17 ICSNNSGKDGNTSANSADSVKGNLTLEISKKITDSNAVLLAVKEVALLSIDEIATAKA 76
DB 61 IGGKIHONNGLDTEENNHNHNSGLAGAYAISTLIKOKLDGKLNKKEKIDAKKCEFTFTN 120
DB 77 IGGKIHONNGLDTEENNHNHNSGLAGAYAISTLIKOKLDGKLNKKEKIDAKKCEFTFTN 136
QY 121 KLEKHTDLDKEGVTDADAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 137 KLEKHTDLDKEGVTDADAEALITKNGTKGAEELGKLFESVEVLSKAKEMLANSVK 196
QY 181 ELTS 184
DB 197 ELTS 200

RESULT 6
Q9R7B2 PRELIMINARY: PRT: 182 AA.
AC Q9R7B2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX NCBI_TaxID=139;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TESTS;
RX MEDLINE=97478003; Pubmed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?"
  Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RL
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TESTS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB81894.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match 47.6%; Score 885; DB 2; Length 182;
Best Local Similarity 98.9%; Pred. No. 9,1e-34;
Matches 180; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 GKDGTANSANSADSVKGNLTLEINKKITDSNAVLLAVKEVALLSIDEIATAKAIGKKIH 66
DB 1 GKDGTANSANSADSVKGNLTLEISKKITDSNAVLLAVKEVALLSIDEIATAKAIGKKIH 60
QY 67 QNNGCLDTEENNHNHNSGLAGAYAISTLIKOKLDGKLNKKEKIDAKKCEFTFTN 126
DB 61 QNNGCLDTEENNHNHNSGLAGAYAISTLIKOKLDGKLNKKEKIDAKKCEFTFTN 120
QY 127 TDLDKEGVTDADAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVKELTSPV 186
DB 121 TDLDKEGVTDADAEALITKNGTKGAEELGKLFESVEVLSKAKEMLANSVKELTSPV 180
QY 187 VA 188
DB 181 VA 182

RESULT 7
Q31120 PRELIMINARY: PRT: 191 AA.
AC Q31120;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS PC.
DB Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=OC10;
RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF029869; AAB86552.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER
SQ SEQUENCE 191 AA; 20502 MW; E4FB56BCB61740F9 CRC64;

Query Match 47.6%; Score 885; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 9,6e-34;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 198 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257  
 DB 10 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 69  
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 317  
 DB 70 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 129  
 QY 318 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 377  
 DB 130 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 189  
 QY 378 S 378  
 DB 190 S 190

LT 8  
 O953P2 PRELIMINARY: PRT; 192 AA.  
 AC O953P2: PRELIMINARY: PRT; 192 AA.  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 GN OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=297;  
 RX MEDLINE=96296448; PubMed=8709845;  
 RA Livey I., Gibbs C.P., Schuster R., Dorner F.,  
 RT "Evidence for lateral transfer and recombination in OspC variation in  
 RL Lyme disease Borrelia."  
 DR Mol. Microbiol. 18:257-269(1995).  
 DR EMBL: L42893; AAB37001.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT 192  
 SO SEQUENCE 192 AA; 20472 MW; 46AC8F93E4DFED6C CRC64;

Query Match 47.3%; Score 879; DB 2; Length 192;  
 Best Local Similarity 99.4%; Pred. No. 1,8e-33;  
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257  
 DB 2 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 61  
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 317  
 DB 62 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 121  
 QY 318 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 377  
 DB 122 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 181  
 QY 378 S 378  
 DB 182 S 182

RESULT 9  
 ID O44719 PRELIMINARY: PRT; 210 AA.  
 AC O44719;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE OSPC.  
 GN Borrelia burgdorferi (Lyme disease spirochete).  
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HB19;  
 RX MEDLINE=95154673; PubMed=7851744;  
 RA Stevenson B., Barthold S.W.;  
 RT "Expression and sequence of outer surface protein C among North  
 RL American isolates of Borrelia burgdorferi."  
 DR EMBL: U04281; AAC43297.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 SO SEQUENCE 210 AA; 22526 MW; 2D672991D584E4EE CRC64;

Query Match 47.3%; Score 879; DB 2; Length 210;  
 Best Local Similarity 99.4%; Pred. No. 2e-33;  
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257  
 DB 20 NNSGDKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 79  
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 317  
 DB 80 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCSSEFTAKLK 139  
 QY 318 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 377  
 DB 140 GEHTDLGKEGYTDDNAKKAIIKTNNDKTKGADELEKLFESVKNLSKAAREMLTNSVKELT 199  
 QY 378 S 378  
 DB 200 S 200

RESULT 10  
 ID O44999 PRELIMINARY: PRT; 177 AA.  
 AC O44999;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 GN OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TXGW;  
 RX MEDLINE=95286481; PubMed=7768799;  
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC."  
 DR EMBL: X84783; CA53234.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT 177  
 SO SEQUENCE 177 AA; 18800 MW; 156671B9614E7A2D CRC64;

Query Match 45.9%; Score 853; DB 2; Length 177;



Best Local Similarity 98.3%; Pred. No. 2,5e-32;  
Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNG 70  
|||||  
DB 1 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNG 60  
|||||

QY 71 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFNKLEKHTDIDG 130  
|||||  
DB 61 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFNKLEKHTDIDG 120  
|||||

QY 131 KEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSKVETLSPVY 187  
| |||||  
DB 121 KEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSKVETLSPVY 177  
| |||||

RESULT 11  
Q9REH7 PRELIMINARY; PRT: 178 AA.

AC Q9REH7 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
GN OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
NCBI\_TaxID=139;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-B31.  
RX MEDLINE-94075528; PubMed-8233951;  
RA Theisen M., Frederiksen B., Ledech A.M., Vuust J., Hansen K.;  
RT "Polymorphism in ospc gene of Borrelia burgdorferi and  
RT immunoreactivity of Ospc protein: Implications for taxonomy and for  
RT use of Ospc protein as a diagnostic antigen.";  
RL J. Clin. Microbiol. 31:2570-2576(1993).  
DR EMBL: X73622; CAA52001.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1 178  
FT SEQUENCE 178 AA; 18894 MW; D619AE6C646F14EB CRC64;

Query Match 45.4%; Score 843.5; DB 2; Length 178;  
Best Local Similarity 97.8%; Pred. No. 6,8e-32;  
Matches 174; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

DB 11 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNG 70  
|||||  
DB 1 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNG 60  
|||||

QY 71 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFNKLEKHTDIDG 129  
|||||  
DB 61 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFNKLEKHTDIDG 120  
|||||

QY 130 KEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSKVETLSPVY 187  
| |||||  
DB 121 KEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSKVETLSPVY 178  
| |||||

RESULT 12  
Q9R7B4 PRELIMINARY; PRT: 175 AA.

AC Q9R7B4 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
GN OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TETS;  
RX MEDLINE-97478003; PubMed-9336916;  
RA Ras N.M., Postic D., Foretz M., Baranton G.;  
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
RT U.S.A.'";  
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TETS;  
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U91792; AAB81889.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1 175  
FT SEQUENCE 175 AA; 18573 MW; EC059E7BD3AC3250 CRC64;

Query Match 44.6%; Score 829; DB 2; Length 175;  
Best Local Similarity 97.6%; Pred. No. 3,1e-31;  
Matches 166; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAKA 60  
| |||||  
DB 6 ISCNNSGKDGNTSANSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAKA 65  
| |||||

QY 61 ICKKHQNGGLDPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFN 120  
|||||  
DB 66 ICKKHQNGGLDPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAACCSEFTFN 125  
|||||

QY 121 KLEKHTDIDGKEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKA 170  
| |||||  
DB 126 KLEKHTDIDGKEGVTDADAKAAILKANGTKTGAEELGKLFESVEVLSKA 175  
| |||||

RESULT 13  
Q9R7A9 PRELIMINARY; PRT: 163 AA.

AC Q9R7A9 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
GN OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
NCBI\_TaxID=139;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TETS;  
RX MEDLINE-97478003; PubMed-9336916;  
RA Ras N.M., Postic D., Foretz M., Baranton G.;  
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
RT U.S.A.'";  
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TETS;  
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U91801; AAB81898.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1 163  
FT SEQUENCE 163 AA; 17210 MW; 7550141651BD01FF CRC64;

Query Match 43.1%; Score 802; DB 2; Length 163;  
 Best Local Similarity 98.2%; Pred. No. 4.8e-30;  
 Matches 160; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVYKGNLTKITDSNNAVLAVKEVALLSIDEIAAKAI 61  
 DB 1 SCNNSGKDGNTSANSADSVYKGNLTKITDSNNAVLAVKEVALLSIDEIAAKAI 60

QY 62 GKRIHONNGIDTENNHNHSLAGAVASTLIKOKLDGKNEGLKEKIDAKKCEFTTNK 121  
 DB 61 GKRIHONNGIDTENNHNHSLAGAVASTLIKOKLDGKNEGLKEKIDAKKCEFTTNK 120

QY 122 LKEKHTDLGKEGVTDADAKKAILKANGTKTGAEELKLFESV 164  
 DB 121 LKEKHTDLGKEGVTDADAKKAILKANGTKTGAEELKLFESV 163

LT 14  
 ID 09AGBI PRELIMINARY; PRT; 165 AA.  
 AC 09AGBI;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OSPC-MODIFIED.  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HB19;  
 RA Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,  
 RA Lawson C.L., Swaminathan S.;  
 RT "Crystal structure of outer surface protein C (OspC) from the Lyme  
 RT disease spirochete, Borrelia burgdorferi.";  
 RL EMBL J. 0:0-0(2001).  
 DR EMBL: AF37548; AAK21289.1; M -> I.  
 FT VARIANT 61 61  
 SQ SEQUENCE 165 AA; 17833 MW; 56E01536D22F61BF CRC64;

Query Match 42.3%; Score 786; DB 2; Length 165;  
 Best Local Similarity 98.2%; Pred. No. 2.6e-29;  
 Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 215 VGPMLTEISKITTSNAVLAVKEVETLTSIDELAKAIGKRIKNDVSLDNEADHNSL 274  
 DB 1 MKGPNLTKITTSNAVLAVKEVETLTSIDELAKAIGKRIKNDVSLDNEADHNSL 60

QY 275 ISGALISMLTKKISAKDSGELKAIEKAKKCEFTTAKIKGHTDLGKEGVTDNAK 334  
 DB 61 MSGAVLISTLTKKISAKDSGELKAIEKAKKCEFTTAKIKGHTDLGKEGVTDNAK 120

QY 335 KAILTNDKTKGADELKLFESVKNLSKAKEMLTNSVKELTS 378  
 DB 121 KAILTNDKTKGADELKLFESVKNLSKAKEMLTNSVKELTS 164

RESULT 15  
 ID 044720 PRELIMINARY; PRT; 211 AA.  
 AC 044720;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OSPC.  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=25015;  
 RX MEDLINE=95154673; PubMed=7851744;  
 RA Stevenson B., Barthold S.W.;  
 RT "Expression and sequence of outer surface protein C among North  
 RT American isolates of Borrelia burgdorferi.";  
 RL FEMS Microbiol. Lett 124:367-372(1994).  
 DR EMBL: U04282; AAC5540.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 SQ SEQUENCE 211 AA; 22603 MW; 63984BA6D8743ED5 CRC64;

Query Match 41.7%; Score 774.5; DB 2; Length 211;  
 Best Local Similarity 81.0%; Pred. No. 1.1e-28;  
 Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 1 MACNNSGKDGNTSANSADSVYKGNLTKITDSNNAVLAVKEVALLSIDEIAAK 59  
 DB 17 ISCNNSGKDGNTSANSADSVYKGNLTKITDSNNAVLAVKEVALLSIDEIAAK 76

QY 60 AIGKRIHONNGIDTENNHNHSLAGAVASTLIKOKLDGKNEGLKEKIDAKKCEFTT 119  
 DB 77 AIGKRIHONNGIDTENNHNHSLAGAVASTLIKOKLDGKNEGLKEKIDAKKCEFTT 136

QY 120 NKLKHTDLGKEGVTDADAKKAILKANGTKTGAEELKLFESVYLSKAKEMLANSV 179  
 DB 137 NKLKSHTELGKQDDDDAKKAILRTKDKGAELDKLFKAVENLSKAKEMLANSV 196

QY 180 KELTSPVVAESPKPK 194  
 DB 197 KELTSPVVAESPKPK 211

Search completed: March 18, 2002, 10:10:53  
 Job time: 976 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:30 ; Search time 118.14 Seconds

(without alignments)  
236.377 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850

Sequence: 1 MACNNSGKDCGNTSANSADSE.....AVENLAKAKEMIANSVKEL 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.1101.\*

1: /SID8/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1850	100.0	377	22 AAB62713
2	1825	98.6	400	22 AAB62739
3	1796.5	97.1	373	22 AAB62711
4	1793.5	96.9	397	22 AAB62729
5	1614.5	87.3	378	22 AAB62712
6	1589.5	85.9	401	22 AAB62738
7	1561	84.4	374	22 AAB62710
8	1558	84.2	398	22 AAB62728
9	1519.5	82.1	369	22 AAB62716
10	1516.5	82.0	393	22 AAB62731
11	1480.5	80.0	367	22 AAB62719

12	1477.5	79.9	391	22 AAB62735	Borrelia sp chimera
13	1467	79.3	384	22 AAB62726	Borrelia sp chimera
14	1465	79.2	560	22 AAB62724	Borrelia sp chimera
15	1464	79.1	408	22 AAB62737	Borrelia sp chimera
16	1284	69.4	370	22 AAB62715	Borrelia sp chimera
17	1281	69.2	394	22 AAB62730	Borrelia sp chimera
18	1259.5	68.1	369	22 AAB62714	Borrelia sp chimera
19	1256.5	67.9	393	22 AAB62732	Borrelia sp chimera
20	1245	67.3	368	22 AAB62718	Borrelia sp chimera
21	1242	67.1	392	22 AAB62734	Borrelia sp chimera
22	1225.5	66.2	368	22 AAB62736	Borrelia sp chimera
23	1217.5	65.8	391	22 AAB62735	Borrelia sp chimera
24	1183.5	64.0	378	22 AAB62725	Borrelia sp chimera
25	1173.5	63.5	401	22 AAB62733	Borrelia sp chimera
26	1164	62.9	410	22 AAB62740	Borrelia sp chimera
27	1163	62.9	386	22 AAB62727	Borrelia sp chimera
28	972.5	52.6	466	16 AAR75740	B31 outer surface
29	954	51.6	209	22 AAB62720	B burgdorferi ospc
30	945	51.1	210	16 AAM11935	Outer surface prot
31	943	51.0	192	15 AAR60886	Borrelia IP2 OSpC
32	939	50.8	192	22 AAB62721	B burgdorferi ospc
33	929	50.2	587	16 AAR75746	B31 Osp-A/antigen
34	928	50.2	210	16 AAR75727	B burgdorferi str
35	915	49.5	466	16 AAR75739	B31 outer surface
36	908	49.1	192	22 AAB62703	B burgdorferi ospc
37	876	47.4	193	22 AAB62709	B burgdorferi ospc
38	864	46.7	176	15 AAR62772	Borrelia B31 anti
39	707	38.2	192	15 AAR60889	Borrelia 297 anti
40	694.5	37.5	177	15 AAR62773	Borrelia 25015 ant
41	694.5	37.5	177	15 AAR60887	Borrelia 25015 Osp
42	692.5	37.4	211	18 AAM11822	B burgdorferi sen
43	692.5	37.4	211	20 AAM93473	B burgdorferi Osp
44	690.5	37.3	193	15 AAR60888	Borrelia Z57 anti
45	685	37.0	212	16 AAM11934	Outer surface prot

## ALIGNMENTS

RESULT 1

1 AAB62713 standard; Protein; 377 AA.

XX AAB62713;

XX 03-APR-2001 (first entry)

DE Borrelia sp chimera ospc protein SEQ ID NO: 30.

XX Borrelia; ospc; Lyme disease; vaccine; chimera protein; tick.

XX Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000MO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNT ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattayler RJ, Selnost G, Dykhuzen D, Luft BJ, Gomes-Solbeck M;

XX WPI: 2001-050113/06.

XX N-PSDB: AAR29017.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

XX to Lyme disease -

XX Claim 43; Page 86-87; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 377 AA;

Query Match 100.0%; Score 1850; DB 22; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-124;  
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKRTDSNVLAVKEVALLSSIDEIAKA 60

OY 1 macnsgkdgntsansadesvkgpnlteinkrtidsnavlavkeveallssideiaaka 60

OY 61 IGKRTIHONNGLDTENNHNGSLLAGAVAI STLKOKLDGLKNGLEKERTIDAKKCEFTFN 120

OY 61 igkrtihngngldteennhngsllagayaistlikqldglneglekeidaakkcetftn 120

OY 121 KLEKHTDGLKEGYTDADAKPAIKANGTKTGAELGKLFESVEVLSKAKEMLANSVK 180

OY 121 klekhtdglkegytdadakeaalkangtktkgaeeigklfesevylskaakemlansvk 180

OY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRTTESNAVLA VKET 240

OY 181 eltspvvaespkkpsmvnnsqgdgntsansadesvkgpnlteiskrttesnavlavk 240

OY 241 ETLASIDELATKATGKTIQONGSLAVEAGHNGTLA GAVTISKLTITOKLDGLKNSK 300

OY 241 etllasidelatkaigkktiqngslaveagngtllagaytiskltitqldglnsek 300

OY 301 EKIEKNAKCEDEFTKLEGEHAQIGIENVTDENAKKAILITDAKKGAAELKLFKAVE 360

OY 301 ekienakcsedftkkllegehaqigienvtdenakkaillitdaakkgaaeklfkave 360

OY 361 NLAKAKEMLANSVKEL 377

OY 361 nlakaakemlansvkel 377

RESULT 2  
 ID AAB62739 standard; Protein: 400 AA.

AC AAB62739;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dettwyler RJ, Sehnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29043.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

PT to Lyme disease -

XX Claim 43; Page 153; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and

CC chimeric ospc proteins from members of the Borrelia genus. These may be

CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as

CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 400 AA;

Query Match 98.6%; Score 1825; DB 22; Length 400;  
 Best Local Similarity 98.9%; Pred. No. 1,1e-122;  
 Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNVLAVKEVALLSSIDEIAKA 61

OY 25 scnsgkdgntsansadesvkgpnlteiskrtidsnavlavkeveallssideiaaka 84

OY 62 GKRTIHONNGLDTENNHNGSLLAGAVAI STLKOKLDGLKNGLEKERTIDAKKCEFTFN 121

OY 85 gkrtihngngldteennhngsllagayaistlikqldglneglekeidaakkcetftn 144

OY 122 LKEKHTDGLKEGYTDADAKPAIKANGTKTGAELGKLFESVEVLSKAKEMLANSVK 181

OY 145 lkekhtdglkegytdadakeaalkangtktkgaeeigklfesevylskaakemlansvk 204

OY 162 LTSPPVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRTTESNAVLA VKET 241

OY 205 ltsppvvaespkkpsmvnnsqgdgntsansadesvkgpnlteiskrttesnavlavk 264

OY 242 TLLASIDELATKATGKTIQONGSLAVEAGHNGTLA GAVTISKLTITOKLDGLKNSK 301

OY 265 tllasidelatkaigkktiqngslaveagngtllagaytiskltitqldglnsek 324

OY 302 KIENAKKCEDEFTKLEGEHAQIGIENVTDENAKKAILITDAKKGAAELKLFKAVE 361

OY 325 kienakcsedftkkllegehaqigienvtdenakkaillitdaakkgaaeklfkave 384

OY 362 LAKAKEMLANSVKEL 377

OY 385 lakaakemlansvkel 400

RESULT 3  
 ID AAB62711 standard; Protein: 373 AA.

AC AAB62711;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 26.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN W0200078966-A1.  
 PD 28-DEC-2000.  
 PF 19-JUN-2000; 2000WO-US16915.  
 PR 18-JUN-1999; 99US-0140042.

```

XX PA (UNIV ) UNIV NEW YORK STATE RES FOUND.
PA PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Dattwyler RJ, Selnost G, Dykhuizen D, Lwift BJ, Gomes-Solecki M;
XX WP1: 2001-050113/706.
DR N-PSDB; AAF29015.
XX
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX Claim 43; Page 81; 160pp; English.
PS
XX The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrella burgdorferi, B. azetelli or B. garinii. These can be used as
CC vaccines against Borrellia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 373 AA:
XX
Query Match 97.1%; Score 1796.5; DB 22; Length 373;
Best Local Similarity 98.1%; Pred. No. 1e-120;
Matches 369; Conservative 2; Mismatches 2; Indels 3; Gaps 1.
OY 2 ACNNGSKDGNFSANSADSVKGPNTLETINKKITDSNAVLAVKEVEALLSIDETAAKAI 61
Db 1 acnngskdgnfsansadesvkgpnlteliskkitdsnavllavkeveallssidelaakai 60
OY 62 GKIKINQNNGLDTENNHNHNSGLAGAVAISTLIKOKLDGKLNKGELKEKIDAAKKCSETFTNK 121
Db 61 gkikinqnmgldeyghngsllagayajserlltkqridglneglkexidaakkcssetftnk 120
OY 122 LKEKITDLGKESVTDADAKEAIIKANGTKTGCAEELGKFESVEVLSKAAKEMLANSVKE 181
Db 121 lkekhtldlgkevtdadakeaillktngctckgaeeigkfesvevlaskaakemlansvke 180
OY 182 LTSPVAESPKRPMNVNNSGKDGNFSANSADSVKGPNTLETISKITTESNAVVLAVKEIE 241
Db 181 ltspvaas--pramvnnsqkdgnfsansadesvkgpnlteliskkitesnnavvlavkeie 237
OY 242 TLLASIDELATAIKAIGKKIQOONGSLAVEAGHGNTLLAGAYTISKLITQKLDGKNSEKLE 301
Db 238 tllasidelataikaigkkiqngslaveaghgntllagaytisklitqrldgknsekle 297
OY 302 KIENAKKCSDEFTFKLCEGHAQLGIENVTDENAKKAILTTDAKDKGAAELEKRAVEN 361
Db 298 kienakkcsedctfklegehaqlgienvtdenakkailttdaakdkgaeleklfxaven 357
OY 362 LAKAKEMLANSVKEL 377
Db 358 lakaakemlansvkcl 373
XX
RESULT 4
AAB62729
ID AAB62729 standard; Protein; 397 AA.
XX
XX AAB62729;
AC
XX
XX 03-APR-2001 (first entry)
DT
XX
XX Borrella sp chimeric ospc protein SEQ ID NO: 62.
KW Borrella; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrella sp.
OS Chimeric - Borrella sp.
XX
PN WO200078966-AI.
```

XX	28-DEC-2000.
PD	
XX	19-JUN-2000; 2000WO-US16915.
PF	
XX	18-JUN-1999; 99US-0140042.
PR	
XX	(UNY ) UNIV NEW YORK STATE RES FOUND.
PA	(BROO-) BROOK BIOTECHNOLOGIES INC.
XX	
PI	Dattwyler RJ, Seinoest G, Dykhulzen D, Luft BJ, Gomes-Solecki M;
XX	
DR	WPI: 2001-050113/06.
XX	N-PSDB: AAF29033.
PT	Compositions of Ospc polypeptides from strains of Borrelia which cause
PT	Lyme disease are used to immunize animals and detect immune responses
PT	to Lyme disease -
PS	
XX	Claim 43; Page 125-126; 160pp; English.
XX	
CC	The present invention provides compositions comprising ospc proteins and
CC	chimeric ospc proteins from members of the Borrelia genus. These may be
CC	Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC	vaccines against Borrelia infection, which is spread by ticks and leads
CC	to Lyme disease.
XX	
SQ	Sequence 397 AA;
XX	
Query Match	96.9%; Score 1793.5; DB 22; Length 397;
Best Local Similarity	97.9%; Pred. No. 1.9e-120;
Matches 368; Conservative	3; Mismatches 2; Indels 3; Gaps 1.
OY	2 ACNNGSKDGMTSANSADSVKGNLTETNKRTTDSNAVYLVAKVEYALSTIDEIAAKAI 61
DB	25 scnnsgkdgnitsansadesvkgpnlteisklidsnavllavkeveallssideiaakai 84
OY	62 GKIKHONNGLDITENNHNHNSILGAVAIISTLIKOKLDGLNEGLKEKIDAKKCSFTYTK 121
DB	85 gkikngmgldtegyhngsllagayalstlltkqldglnegllkeidakkcsftnk 144
OY	122 LKEKHTDLGKEGVTADAKAEALILKANGTKTGAABELGKLFESVEVLSKAAKEMLANSVKE 181
DB	145 lkehtldlgkegvtdadakeailltkngctkgaeeigklfesvevlskaakemlansvke 204
OY	182 LTPSYVASEPRKPSRVNNGSKDGNITSANSADSVKGNLTETNKRTTDSNAVYLVAKVE 241
DB	205 ltpsyvae--pawvnnsgkdgnitsansadesvkgpnlteisklitesnavllavkele 261
OY	242 TLIASIDELATAIKGKKITQONGGLAVEAGHNTLITAGATTISKLITQOKIDGNKSEKLE 301
DB	262 tllasidelataikgkikqngglaveaghnctllagaytisklltqkldglnesklike 321
OY	302 KLENAKKCSSEDTFKLEGEHQAOLGIEVNTDENAKKAILITDAKDKGAEELEKRAVEN 361
DB	322 klenakkcsedctfkklgegheqglenvtdenakkailitdaakdkgaeelekfraven 381
OY	362 LAKAKEMLANSVKEL 377
DB	382 lakaakemlansvkel 397
RESULT	5
ID	AAB62712
XX	AAB62712 standard; Protein; 378 AA.
XX	
AC	AAB62712;
XX	
DT	03-APR-2001 (first entry)
XX	
DE	Borrelia sp chimeric ospc protein SEQ ID NO: 28.
XX	

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX WO200078966-A1.  
 PN 28-DEC-2000.  
 PD 19-JUN-2000; 2000MO-US16915.  
 PF 18-JUN-1999; 99US-0140042.  
 PR (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 PI WPI: 2001-050113/06.  
 N-PSDB; AAF29016.  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -  
 XX Claim 43; Page 83-84; 160pp; English.  
 PS The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX Sequence 378 AA;  
 SQ

Query Match 87.3%; Score 1614.5; DB 22; Length 378;  
 Best Local Similarity 87.8%; Pred. No. 1.1e-107;  
 Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;  
 QY 1 MACNSGKDGNTSANSADSVKGPNTLEINRKTTSNAVLAVKEVEALLSIDETAKA 60  
 DB 1 macnsgkdntsansadesvkgpnltelntktdenavllavkeveallsideaaka 60  
 QY 61 IGRKHONNGLDTENNHNHNSILAGAVASTLIKOKIDGKNEGKLEKIDAAKCSSEFTN 120  
 DB 61 igkkihqngldtenhngsllagayastlikqkldgkneglekidaakkcsseftn 120  
 QY 121 KLEKHTDGLGEGVTDADAKKAILKANGTKTKGAEELKLFESVEVLSKAKEMLANSVK 180  
 DB 121 klekhtdylgegvtdadaakeailkangtlktgaeelglfesevevlskaakemlansvk 180  
 QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEI 240  
 DB 181 eltspvvaesppkpsmvnnsdgkntsansadesvkgpnltelskttessnavlavkev 240  
 QY 241 ELLLSIDELATKAIGKRIQONGGLAVAGHNGTLGAVTISKLITOKIDGKNSKRLK 300  
 DB 241 elllsidela-kalgkrikndvsldeahngsllsagaylislitlkkksaidsqelk 299  
 QY 301 EKIEAKKCSSEFTKLEGEHAOLGIENVTDENAKKAILTTDAKKDGALEKTRFAVE 360  
 DB 300 aeleakkcseefaklgehtdylgegvtdadnakkailktndktkgadelaklesvyk 359  
 QY 361 NLAAKAKMLANSVKEL 377  
 DB 360 nlskaakemltnsvkel 376

RESULT 6  
 ID AAB62738  
 XX AAB62738 standard; Protein; 401 AA.

AC AAB62738;  
 XX 03-APR-2001 (first entry)  
 DT Borrelia sp chimeric ospc protein SEQ ID NO: 80.  
 XX DE Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX OS Chimeric - Borrelia sp.  
 XX OS Chimeric - Borrelia sp.  
 XX WO200078966-A1.  
 PN 28-DEC-2000.  
 PD 19-JUN-2000; 2000MO-US16915.  
 PF 18-JUN-1999; 99US-0140042.  
 PR (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 PI WPI: 2001-050113/06.  
 N-PSDB; AAF29042.  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -  
 XX Claim 43; Page 150-151; 160pp; English.  
 PS The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX Sequence 401 AA;  
 SQ

Query Match 85.9%; Score 1589.5; DB 22; Length 401;  
 Best Local Similarity 86.7%; Pred. No. 7.2e-106;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;  
 QY 2 ACNNSGKDGNTSANSADSVKGPNTLEINRKTTSNAVLAVKEVEALLSIDETAKA 61  
 DB 2 acnnskgdntsansadesvkgpnltelstkttdenavllavkeveallsideaaka 61  
 QY 62 GKKIHONNGLDTENNHNHNSILAGAVASTLIKOKIDGKNEGKLEKIDAAKCSSEFTN 121  
 DB 62 gkkihqngldtenhngsllagayastlikqkldgkneglekidaakkcsseftn 121  
 QY 122 LKEKHTDGLGEGVTDADAKKAILKANGTKTKGAEELKLFESVEVLSKAKEMLANSVK 181  
 DB 122 lkekhtdylgegvtdadaakeailkangtlktgaeelglfesevevlskaakemlansvk 181  
 QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEI 241  
 DB 181 eltspvvaesppkpsmvnnsdgkntsansadesvkgpnltelskttessnavlavkev 241  
 QY 241 ELLLSIDELATKAIGKRIQONGGLAVAGHNGTLGAVTISKLITOKIDGKNSKRLK 301  
 DB 241 elllsidela-kalgkrikndvsldeahngsllsagaylislitlkkksaidsqelk 301  
 QY 301 EKIEAKKCSSEFTKLEGEHAOLGIENVTDENAKKAILTTDAKKDGALEKTRFAVE 361  
 DB 302 aeleakkcseefaklgehtdylgegvtdadnakkailktndktkgadelaklesvyk 363  
 QY 362 LAAKAKMLANSVKEL 377  
 DB 384 lsaakemltnsvkel 399



```
RESULT 7
AAB62710
ID AAB62710 standard; Protein: 374 AA.
XX
AC AAB62710;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 24.
XX
KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000: 2000WO-US16915.
XX
PR 18-JUN-1999: 99US-0140042.
XX
PA (UANY ) UNIV NEW YORK STATE RES FOUND.
XX
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX MPI: 2001-050113/06.
XX
DR N-PSDB: AAF29014.
XX
PT Compositions of ospc polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 78-79; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
XX
XX chimeric ospc proteins from members of the Borrelia genus. These may be
XX
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX
XX to Lyme disease.
XX
SQ Sequence 374 AA;

Query Match 84.4%; Score 1561; DB 22; Length 374;
Best Local Similarity 85.9%; Pred. No. 7.1e-104;
Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSKDGNTSANSADSVKGPNTLTKKIRKIDSNVAVLAVKEVALLSSIDEIAAKAI 61
DB 1 acnsgkdgnstansadesvkgpnltelstktidsnavllavkevellssideiaakal 60
OY 62 GKRIHQNGIDTENNNGSLAGAVYASTLIKOKLDGLKNEGKLEKIDAAKCSSETFNK 121
DB 61 gkkihngngidteyngngslagayaistlikqldglnegklekidaakcssetfntk 120
OY 122 LKKEHNDLGEGVTDADAKFAILKANGTKTKGAEEGLKFESEVVISKAKEMLANSVKE 181
DB 121 lkekhdldggyvtdadakaailktnglctkgaeeigkltesevviskaakemlansvke 180
OY 182 LTPVVAESPKKPSMVNNSKDGNTSANSADSVKGPNTLTKKIRKIDSNVAVLAVKEIE 241
DB 181 ltpvvaes---pamvnnsgkdgnstansadesvkgpnltelstktidsnavllavkeve 237
OY 242 TLLASIDELATKAIGKKIQONGSLAVAGNGTFLAGAVYISTLIKTOKLDGLKNSKLEKE 301
DB 238 tlltsidela-kaiqgkikndvidneadngslisgayslisslntlekkisaiksgelka 296
OY 302 KIENAKKCSDFTKKLEGEHAOLGIENVYVDENAKKAILITIDAAKDGALEKLFKRVEN 361
```

```
DB 297 etekakkcsseftaklkgendldgkegyvtdadnakallkcnoktktgadalekflfsvkn 356
OY 362 LAKAKEMLANSVKEI 377
DB 357 lskakemlansvkel 372

RESULT 8
AAB62728
ID AAB62728 standard; Protein: 398 AA.
XX
AC AAB62728;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 60.
XX
KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000: 2000WO-US16915.
XX
PR 18-JUN-1999: 99US-0140042.
XX
PA (UANY ) UNIV NEW YORK STATE RES FOUND.
XX
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX MPI: 2001-050113/06.
XX
DR N-PSDB: AAF29032.
XX
PT Compositions of ospc polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 123; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
XX
XX chimeric ospc proteins from members of the Borrelia genus. These may be
XX
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX
XX to Lyme disease.
XX
SQ Sequence 398 AA;

Query Match 84.2%; Score 1558; DB 22; Length 398;
Best Local Similarity 85.6%; Pred. No. 1.3e-103;
Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSKDGNTSANSADSVKGPNTLTKKIRKIDSNVAVLAVKEVALLSSIDEIAAKAI 61
DB 25 acnsgkdgnstansadesvkgpnltelstktidsnavllavkevellssideiaakal 84
OY 62 GKRIHQNGIDTENNNGSLAGAVYASTLIKOKLDGLKNEGKLEKIDAAKCSSETFNK 121
DB 85 gkkihngngidteyngngslagayaistlikqldglnegklekidaakcssetfntk 144
OY 122 LKKEHNDLGEGVTDADAKFAILKANGTKTKGAEEGLKFESEVVISKAKEMLANSVKE 181
DB 121 lkekhdldggyvtdadakaailktnglctkgaeeigkltesevviskaakemlansvke 204
OY 145 lkekhdldggyvtdadakaailktnglctkgaeeigkltesevviskaakemlansvke 204
OY 182 LTPVVAESPKKPSMVNNSKDGNTSANSADSVKGPNTLTKKIRKIDSNVAVLAVKEIE 241
DB 205 ltpvvaes---pamvnnsgkdgnstansadesvkgpnltelstktidsnavllavkeve 261
```

```

QY 242 TLASIDELATKAIGKTIQONGGLAVEAGHNGTLTAGAVTISKLTITQKLDGKNSSEKLE 301
DB 262 TLITSLIDEIA-kaigkiknovslidneadngslisgailslitkksaikkdsqelk 320
QY 302 KIENAKKSEDFPKKLEGEHAOLGIENVTDENAKKALLITDAKDKGALEKLFKAVEN 361
DB 321 etekakkseeftaklgehtldlgkevtdnakkalkitnndkttgadelcklfsesvk 380
QY 362 LAKAKEMLANSVKEL 377
DB 381 lskakemltnsvkel 396

RESULT 9
AAB62716
ID AAB62716 standard; Protein; 369 AA.
AAB62716:
DT 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 36.
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX PD 28-DEC-2000.
XX PF 19-JUN-2000; 2000WO-US16915.
XX PR 18-JUN-1999; 99US-0140042.
PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Datwyler RJ, Selinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX DR WPI; 2001-050113/06.
XX DR N-PSDB; AAF29020.
XX PF Compositions of Ospc polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PT to Lyme disease -
PS Claim 43; Page 94-95; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX CC
XX Sequence 369 AA:
QY Query Match 82.1%; Score 1519.5; DB 22; Length 369;
Best Local Similarity 86.2%; Pred. No. 6.5e-101;
Matches 326; Conservative 10; Mismatches 31; Indels 11; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNNAVLAKEVEALLSSIDEIAAKAI 61
DB 1 acnnsykgntsansadesvkgpnlteiskltidsnavllavkeveallssidel-akai 59
QY 62 GKTIHONGGLTENNHNHSLAGAYSTLTKQKLDGK-NEGKKEKIDAAKCKSEFTFN 120
DB 60 gkklndgslneanrnesllagaystltltqklsklnsgelkkaakckseefst 119
QY 121 KLEKHTDGLKEGVTDADAKPAILKANGT-KTKGAELGKLFESVEVLAKAKEMLANSV 179

```

```

DB 120 kldnhaqligvtdenakkalkkanaagkdkgyeeleklsjleslskaakemlansv 179
QY 180 KELTSPVVIESPKRPSMVNNSGKDGNTSANSADSVKGNLTETINKRTIDSNNAVLAKE 239
DB 180 keltspvvg-----nmsgkdgntsansadesvkgpnlteiskltlesnavlavke 231
QY 240 IETLASIDELATKAIGKTIQONGGLAVEAGHNGTLTAGAVTISKLTITQKLDGKNSSEK 299
DB 232 ietllasidelatkaigkligqngglaveaghnngtltagaytisklticqldglnsekl 291
QY 300 KEKIEAKKSEDFPKKLEGEHAOLGIENVTDENAKKALLITDAKDKGALEKLFKAV 359
DB 292 kekienakkseefcklgeghagligenvtdenakkallitdaakdkgaaleklfkav 351
QY 360 ENLAKAKEMLANSVKEL 377
DB 352 enlakaakemlansvkel 369

RESULT 10
AAB62731
ID AAB62731 standard; Protein; 393 AA.
AAB62731:
AC AAB62731;
DT 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 66.
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX PD 28-DEC-2000.
XX PF 19-JUN-2000; 2000WO-US16915.
XX PR 18-JUN-1999; 99US-0140042.
PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Datwyler RJ, Selinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX DR WPI; 2001-050113/06.
XX DR N-PSDB; AAF29035.
XX PT Compositions of Ospc polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PT to Lyme disease -
PS Claim 43; Page 131; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX CC
XX Sequence 393 AA:
QY Query Match 82.0%; Score 1516.5; DB 22; Length 393;
Best Local Similarity 86.0%; Pred. No. 1.2e-100;
Matches 325; Conservative 11; Mismatches 31; Indels 11; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNNAVLAKEVEALLSSIDEIAAKAI 61
DB 25 acnnsykgntsansadesvkgpnlteiskltidsnavllavkeveallssidel-akai 83

```

QY 62 GKRIHONGDTENNHNGSLAGAVATSTLIKOKLOGLK-NEGKKEKIDAAKCESETFN 120  
 DB 84 gkrlkndgsldneenresllagaytlstlltqklslngseglkexlaakccseest 143  
 QY 121 KLEKHTDLCEGYVDAAKKAAILKANGT-KTGAELGLKFESVEVLSKAEMLANSV 179  
 DB 144 klldnhaqlglgvtidenakkaillkanaagkdkyveeleklsglslesksaakemlansv 203  
 QY 180 KELTSPVVAESPCKPSMVNNSGKDGMTSANSADSVKGPMLTEISKITTESNAVLAKE 239  
 DB 204 keltspvvhg-----nsgkdgnltsansadesvkgpmltelskkltsesnavlavke 255  
 QY 240 IETPLASIDELATKATGKRIQONGGLAVEAGHNGTLLAGAVTTSKLTOKLOGKNSKEXL 299  
 DB 256 Ietllasidelackaigkklqngglaveaghnngtllagaytlstlltqklldgiknsek1 315  
 QY 300 KEKIEKAKCEDFTKLEGEHAOLGIENVYDENAKKAILITDAKDKGAELKLFKAV 359  
 DB 316 kekienakccsedftkkllegehaqlglenvydenakkaillitdaakdkgaaleklfkav 375  
 DB 360 ENLAKAKEMLANSVKEL 377  
 DB 376 enlakaakemlansvkel 393

RESULT 11  
 AAB62719  
 ID AAB62719 standard; Protein: 367 AA.  
 AC AAB62719:  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 42.  
 XX  
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX  
 DR WPI: 2001-050113/06.  
 DR N-PSDB: AAF29023.  
 XX  
 PT Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 102-103; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 Sequence 367 AA:  
 Query Match 80.0%; Score 1480.5; DB 22; Length 367;

Best Local Similarity 82.4%; Pred. No. 3,9e-98;  
 Matches 310; Conservative 23; Mismatches 34; Indels 9; Gaps 2;  
 QY 2 ACNNSGKDGMTSANSADSVKGPMLTEINKITDSNAVLAKEVEALLSIDETIAKAI 61  
 DB 1 acnnskdgnasansadesvkgpmltelskkltsesnavlavkevelldasidelacka1 60  
 QY 62 GKRIHONGDTENNHNGSLAGAVATSTLIKOKLOGKNGLEKIDAAKCESETFNK 121  
 DB 61 gkrlkndgsldneenresllagaytlstlltqklslngseglkexlaakccseest 119  
 QY 122 KLEKHTDLCEGYVDAAKKAAILKANGT-KTGAELGLKFESVEVLSKAEMLANSVKE 181  
 DB 120 kleshavlgldnlddnagralikhanckdgaaleklflavenlakaagdtlknavke 179  
 QY 182 LTSPVVAESPCKPSMVNNSGKDGMTSANSADSVKGPMLTEISKITTESNAVLAKEIE 241  
 DB 180 ltspvvhg-----nsgkdgnltsansadesvkgpmltelskkltsesnavlavke1e 231  
 QY 242 TLLASIDELATKATGKRIQONGGLAVEAGHNGTLLAGAVTTSKLTOKLOGKNSKEXLKE 301  
 DB 232 tllasidelackaigkklqngglaveaghnngtllagaytlstlltqklldgiknsek1ke 291  
 QY 302 KIENAKKCEDFTKLEGEHAOLGIENVYDENAKKAILITDAKDKGAELKLFKAVEN 361  
 DB 292 kienakccsedftkkllegehaqlglenvydenakkaillitdaakdkgaaleklfkaven 351  
 QY 362 LAKAKEMLANSVKEL 377  
 DB 352 lakaakemlansvkel 367

RESULT 12  
 AAB62735  
 ID AAB62735 standard; Protein: 391 AA.  
 AC AAB62735:  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 74.  
 XX  
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX  
 DR WPI: 2001-050113/06.  
 DR N-PSDB: AAF29039.  
 XX  
 PT Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 142; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.  
XX  
SQ Sequence 391 AA:

Query Match 79.9%; Score 1477.5; DB 22; Length 391;  
Best Local Similarity 82.2%; Pred. No. 7e-98;  
Matches 309; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

```

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNAVILAVKEVEALLSSIDEIAAKAI 61
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 25 scmsgkdgntansadesvkgpnlteiskitdsnavilavkeveallsideiaakai 84
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 62 GKIHONNGLDTEENNHNCSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCEFTFNK 121
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 85 gkkl-gngl-aengekntsls-gvaylsdl-iaeklnvklneel-ekidtaqgestetfnk 143
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 122 LKEKHTDGLKEGVTDADAKKATILKANGTKYTGAEELGFESVEVLSKAAKEMLANSVKE 181
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 144 lkehnvlgldltdnaqraillkhanckdgaaleleklfkavenlskaaqdtlknsvke 203
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 182 LTSPVVAESPKKPSKVNNSGKDGNTSANSADSVKGNLTETISKITEESNAVILAVKEIE 241
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 204 ltsplvbg-----nnsqdgntsansadesvkgpnlteiskitdsnavilavkeie 255
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 242 TLASIDELATKAIKGIKIQONGGLAVAGHNGTLLAGAYTISKLTITOKLDGLKNSKELKE 301
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 256 tllasidelakaiqkikqngglaveaghnngtllagaytiskliltqkldglnskelke 315
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 302 KIENAKKCEDEFTKLEGEHAQLGIENVTDENAKKATILTDAAKDGAAELKELFKAVEN 361
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 316 kienakccsedfckliegehaqgienvtdenakkailltdaakdgaaelkelfkaven 375
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 362 LAKAAKEMLANSVKEL 377
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 376 lakaakemlansvkel 391
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

## RESULT 13

AAB62726  
ID AAB62726 standard; Protein; 384 AA.

AC AAB62726;  
XX  
DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospc protein SEQ ID NO: 56.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

PD 19-JUN-2000; 2000WO-US16915.

PF 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

XX N-PSDB; AAF29030.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

PI to Lyme disease -

PS Claim 43; Page 117-118; 160pp; English.  
XX  
CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. atzeli or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX  
SQ Sequence 384 AA:

Query Match 79.3%; Score 1467; DB 22; Length 384;  
Best Local Similarity 81.4%; Pred. No. 3.9e-97;  
Matches 306; Conservative 30; Mismatches 36; Indels 4; Gaps 3;

```

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNAVILAVKEVEALLSSIDEIAAKAI 61
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 62 GKIHONNGLDTEENNHNCSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCEFTFNK 121
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 gkklngngldteyhnngslagayaistlikqkldglnknglkeidaakcetfnk 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 122 LKEKHTDGLKEGVTDADAKKATILKANGTKYTGAEELGFESVEVLSKAAKEMLANSVKE 181
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 121 lkehtdglkegvtdadakeaillkngctkgaeeilgkllfsevevlskaakemlansvke 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 182 LTSPVVAESPKKPSKVNNSGKDGNTSANSADSVKGNLTETISKITEESNAVILAVKEIE 241
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 181 ltspyvaespkpfnngns--gdsasfnpesakgpnlvtiskitdsnavilaveve 238
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 242 TLASIDELATKAIKGIKIQONGGLAVAGHNGTLLAGAYTISKLTITOKLDGLKNSKELKE 301
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 239 allsidel-skaigkikldngtldneanrnesllagayelskiltqlslv1-nseelk 296
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 302 KIENAKKCEDEFTKLEGEHAQLGIENVTDENAKKATILTDAAKDGAAELKELFKAVEN 361
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 297 kikeakdcsgkftcklkshaeigisvqddnakkaillkhtgkdkgaaleelfksles 356
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 362 LAKAAKEMLANSVKEL 377
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 357 lskaagaallnsvkel 372
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

## RESULT 14

AAB62724  
ID AAB62724 standard; Protein; 560 AA.

AC AAB62724;

XX 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospc protein SEQ ID NO: 52.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

PD 19-JUN-2000; 2000WO-US16915.

PF 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29028.

XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -

PS Claim 43; Page 112-113; 160pp; English.

XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 560 AA:

Query Match 79.2%; Score 1465; DB 22; Length 560;  
Best Local Similarity 81.0%; Pred. No. 8.8e-97;  
Matches 306; Conservative 23; Mismatches 45; Indels 4; Gaps 2;

1 MACNNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60  
|||||  
1 macnsgkdgnstansadesvkgpnlteiskitdsnavllavkeveallssideiaaka 60

QY 61 IGGKIHQNNGLDTENNHNHNSLAGAVAIITLTKOKIDGLKNEGKIDAKKCSSEFTFN 120  
|||||

DB 61 Iggkihqnngldteyhnngslagayaislltkqldglnegkidaakkcsseftfn 120

QY 121 KLEKHTDGLGEGVTADAKKEAIIKANGTKTKGAELGKLPESVEVLSKAKEMLANSVK 180  
|||||

DB 121 klekhdtlgyegvtadakeaillkngtktkgaeelgklpesvevlskaakemlansvk 180

QY 181 ELTSPVVAESPFPKPSMVNNSGKDGNT-SANSADSVKGNLTETINKKITESNAVLAKE 239  
|||||

DB 181 eltspvvaes---pangnsgkgydsastnpadesakgpnltetiskitdsnavlavke 237

QY 240 IETLALSIDELATKATIGKRIQONGSLAVAGHNGTLLAAYITISKITOKLDGLKNSKLE 299  
|||||

DB 238 velivisidelaakalqgkldnnglaalnqngslagayaislltelksklklnleel 297

QY 300 KKEIEAKKCSSEFTFKLEGEHAQLGIENVTDENAKKAILITDAKDKGAELKLFKAV 359  
|||||

DB 298 kkeiakakcsseeftnklsgshadlygqdatdhaakailtkhatcdkakefkdlfesy 357

QY 360 ENLAKAKEMLANSVKEL 377  
|||||

DB 358 egllkaagvalnsvkel 375

ULF 15

ID AAB62737 standard; Protein; 408 AA.

AC AAB62737;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 78.

KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000OMO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29041.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -

PS Claim 43; Page 147-148; 160pp; English.

XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 408 AA:

Query Match 79.1%; Score 1464; DB 22; Length 408;  
Best Local Similarity 81.1%; Pred. No. 6.8e-97;  
Matches 305; Conservative 31; Mismatches 36; Indels 4; Gaps 3;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 61  
|||||

DB 25 cnsngkdgnstansadesvkgpnlteiskitdsnavllavkeveallssideiaaka 84

QY 62 GKRIHNNGLDTENNHNHNSLAGAVAIITLTKOKIDGLKNEGKIDAKKCSSEFTFN 121  
|||||

DB 62 gkrihqnngldteyhnngslagayaislltkqldglnegkidaakkcsseftfn 121

QY 85 GKRIHNNGLDTENNHNHNSLAGAVAIITLTKOKIDGLKNEGKIDAKKCSSEFTFN 144  
|||||

DB 85 gkrihqnngldteyhnngslagayaislltkqldglnegkidaakkcsseftfn 144

QY 122 LKEKHTDGLGEGVTADAKKEAIIKANGTKTKGAELGKLPESVEVLSKAKEMLANSVK 181  
|||||

DB 145 lkekhdtlgyegvtadakeaillkngtktkgaeelgklpesvevlskaakemlansvk 204

QY 182 LTPVVAESPFPKPSMVNNSGKDGNTSANSADSVKGNLTETINKKITESNAVLAKE 241  
|||||

DB 205 ltspvvaesfpkpfhngnsy--gdsastnpadesakgpnltetiskitdsnavlavke 262

QY 242 TILASIDELATKATIGKRIQONGSLAVAGHNGTLLAAYITISKITOKLDGLKNSKLE 301  
|||||

DB 263 tilasidelatkaigkriqongslavaghnngtllagayaislltelksklklnleel 320

QY 302 KIENAKKCSSEFTFKLEGEHAQLGIENVTDENAKKAILITDAKDKGAELKLFKAVEN 361  
|||||

DB 321 kienakcsseftfkldshaeligsvqdnnakailtkhngkdgakeleelfxsles 380

QY 362 LAKAKEMLANSVKEL 377  
|||||

DB 381 lsaagaalnsvkel 396

Search completed: March 18, 2002, 09:54:31  
Job time: 329 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds  
(without alignments)  
152.860 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850  
Sequence: 1 MACNNSGKDGNTSANSNADES.....AVENLAKAKEMIANSVKEL 377

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Actual number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	972.5	52.6	466	4	US-08-235-836C-110 Sequence 110, Appl
2	945	51.1	210	1	US-08-158-353-3 Sequence 3, Appl
3	938	50.7	210	4	US-08-209-603E-15 Sequence 15, Appl
4	933	50.4	209	4	US-09-196-293-15 Sequence 15, Appl
5	932	50.4	210	4	US-08-235-836C-30 Sequence 30, Appl
6	929	50.2	588	4	US-08-235-836C-122 Sequence 122, App
7	915	49.5	466	4	US-08-235-836C-107 Sequence 107, App
8	685	37.0	212	1	US-08-158-353-2 Sequence 2, Appl
9	670	36.2	212	1	US-08-031-295-2 Sequence 2, Appl
10	670	36.2	212	1	US-07-903-580-2 Sequence 2, Appl
11	631	34.1	212	1	US-08-158-353-4 Sequence 4, Appl
12	629	34.0	212	4	US-09-196-293-11 Sequence 11, Appl
13	629	34.0	212	4	US-08-209-603E-11 Sequence 11, Appl
14	626	33.8	212	4	US-08-235-836C-34 Sequence 34, Appl
15	614.5	33.2	209	4	US-08-235-836C-32 Sequence 32, Appl
16	592.5	32.0	207	4	US-08-235-836C-36 Sequence 36, Appl
17	171	9.2	194	4	US-09-364-083-2 Sequence 2, Appl
18	168	9.1	630	4	US-08-973-462-9 Sequence 9, Appl
19	164.5	8.9	1164	4	US-08-923-992A-2 Sequence 2, Appl
20	162.5	8.8	1164	4	US-08-923-992A-10 Sequence 10, Appl
21	160	8.5	1786	4	US-08-973-462-8 Sequence 8, Appl
22	158	8.5	1098	4	US-08-923-992A-8 Sequence 8, Appl
23	154.5	8.4	1128	4	US-08-923-992A-6 Sequence 6, Appl
24	151.5	8.2	1104	4	US-08-923-992A-4 Sequence 4, Appl
25	148	8.0	3111	2	US-08-460-309-4 Sequence 4, Appl
26	148	8.0	3111	2	US-08-125-077-4 Sequence 4, Appl
27	147.5	8.0	2482	1	US-08-328-254-6 Sequence 6, Appl

28	146.5	7.9	688	3	US-09-141-047-8 Sequence 8, Appl
29	146.5	7.9	3248	1	US-08-353-700-1 Sequence 1, Appl
30	146.5	7.9	3248	5	PCT-US95-16216-1 Sequence 1, Appl
31	145	7.8	1388	2	US-08-685-576-1 Sequence 1, Appl
32	143.5	7.8	1388	2	US-08-685-576-4 Sequence 4, Appl
33	143	7.7	414	5	PCT-US93-03077-3 Sequence 3, Appl
34	143	7.7	1093	5	PCT-US93-03077-1 Sequence 1, Appl
35	140.5	7.6	1196	1	US-08-144-121-4 Sequence 4, Appl
36	140.5	7.6	1196	2	US-08-735-893-4 Sequence 4, Appl
37	140	7.6	1147	3	US-08-470-260-5 Sequence 5, Appl
38	140	7.6	1147	3	US-08-471-491-5 Sequence 5, Appl
39	140	7.6	1147	4	US-08-466-662-5 Sequence 5, Appl
40	140	7.6	3289	2	US-08-477-451-2 Sequence 2, Appl
41	139.5	7.5	396	1	US-08-430-024-2 Sequence 2, Appl
42	139.5	7.5	396	1	US-08-782-009-2 Sequence 2, Appl
43	139.5	7.5	396	3	US-09-017-302-2 Sequence 2, Appl
44	133.5	7.2	1354	3	US-08-685-871-2 Sequence 2, Appl
45	132	7.1	2285	4	US-09-308-375-2 Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-235-836C-110  
Sequence 110, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: A35  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ. ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 52.6%; Score 972.5; DB 4; Length 466;  
Best Local Similarity 62.9%; Pred. No. 5.8e-66;  
Matches 224; Conservative 23; Mismatches 74; Indels 35; Gaps 8;  
QY 1 MACNNSGKDGNTSANSNADESVPKPNLTETIKKTTDSNAVLAVKEVALLSIDEIAAKA 60

Db 17 ISCNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLAVKEVALLSIDETIAKA 76  
Qy 61 ICKIHHNNGLDTENNHNSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 120  
Db 77 ICKIHHNNGLDTENNHNSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 136  
Qy 121 KLEKHTDGLKEGVTDADAKAELKANGTKTKGAELGKLFESVEVLSKAAREMLANSVK 180  
Db 137 KLEKHTDGLKEGVTDADAKAELKANGTKTKGAELGKLFESVEVLSKAAREMLANSVK 196  
Qy 181 ELTSPVVAESPKPKPQ--NVSSLD--EKNSVVDLPGEKVLVSKERKNGKGY----- 245  
Db 197 ELTSPVVAESPKPKPQ--NVSSLD--EKNSVVDLPGEKVLVSKERKNGKGY----- 245  
Qy 241 ETLASIDELATKAIGKKIQONGGL-AVEAGHNGTLA-----GATIS 283  
Db 246 -DLATVTKLELKGTSMDKNGSGVLEGVKADKSKVKLTISDGLGOTLLEVPKEDGKTLVS 304  
Qy 284 KLITOKLDGLKNSERKLEKINAKKCEDEFTKLEGEHAQ-LGIENTVDENAKKAI 338  
Db 305 KKVTSK-DKSTEEKFNEKE-----VSEKITTRADGTRLEVTGKISDGSKAKEVL 355

RESULT 2  
US-08-158-353-3  
; Sequence 3, Application US/08158353  
; Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158.353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cario, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-3

Query Match 51.1%; Score 945; DB 1; Length 210;  
Best Local Similarity 97.9%; Pred. No. 2.3e-64;  
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MACNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLAVKEVALLSIDETIAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLAVKEVALLSIDETIAKA 76

Qy 61 ICKIHHNNGLDTENNHNSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 120  
Db 77 ICKIHHNNGLDTENNHNSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 136  
Qy 121 KLEKHTDGLKEGVTDADAKAELKANGTKTKGAELGKLFESVEVLSKAAREMLANSVK 180  
Db 137 KLEKHTDGLKEGVTDADAKAELKANGTKTKGAELGKLFESVEVLSKAAREMLANSVK 196  
Qy 181 ELTSPVVAESPKPKPQ--NVSSLD--EKNSVVDLPGEKVLVSKERKNGKGY----- 245  
Db 197 ELTSPVVAESPKPKPQ--NVSSLD--EKNSVVDLPGEKVLVSKERKNGKGY----- 245  
RESULT 3  
US-08-209-603E-15  
; Sequence 15, Application US/08209603E  
; Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILKE, BETTINA  
APPLICANT: PREAC-MORSIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROOKS HAIDT HAEFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCIT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE:  
DESCRIPTION: PROTEIN  
HYPOTHETICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: B. BURGDOFFER  
IMMEDIATE SOURCE:  
LIBRARY: DSM 5662  
POSITION IN GENOME: N/A  
FEATURE:  
IDENTIFICATION METHOD: amino acid analysis  
PUBLICATION INFORMATION: N/A



US-08-209-603E-15

Query Match 50.7%; Score 938; DB 4; Length 210;  
Best Local Similarity 97.4%; Pred. No. 7.9e-64;  
Matches 189; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIATAKA 76  
QY 61 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 136  
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 180  
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 196  
QY 181 ELTSPVAESPCKP 194  
DB 197 ELTSPVAESPCKP 210

RESULT 4

US-09-196-293-15  
Sequence 15, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Sautscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: 738.001US2  
CURRENT APPLICATION NUMBER: US/09/196,293  
CURRENT FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 209  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-15

Query Match 50.4%; Score 933; DB 4; Length 209;  
Best Local Similarity 97.9%; Pred. No. 1.9e-63;  
Matches 188; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIATAKA 76  
QY 61 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 136  
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 180  
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 196

QY 181 ELTSPVAESPCKP 192  
DB 197 ELTSPVAESPCKP 208

RESULT 5

US-08-235-836C-30  
Sequence 30, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Lofit, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Compising  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-30

Query Match 50.4%; Score 932; DB 4; Length 210;  
Best Local Similarity 96.9%; Pred. No. 2.2e-63;  
Matches 188; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAKA 60  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDEIATAKA 76  
QY 61 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 120  
DB 77 ICKKIHNNGLDTENNHNSLAGAAYAISTLIKOKLDGLKNEGLKEKIDAKKCSSETFTN 136  
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 180  
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAEMLANSVK 196  
QY 181 ELTSPVAESPCKP 194  
DB 197 ELTSPVAESPCKP 210

RESULT 6

US-08-235-836C-122  
Sequence 122, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
NUMBER OF INVENTION: Borrelia Polypeptides and Uses Therefor  
CORRESPONDENCE ADDRESS: 144  
STREET: Brookhaven National Laboratory  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 122:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 588 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-122

Query Match 50.2%; Score 929; DB 4; Length 588;  
Best Local Similarity 89.7%; Pred. No. 1.5e-62;  
Matches 191; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

4 NNSGKDGNTSANSADSVKGPNTETNKKITDSNAVLAVKEVEALLSSIDEIAKAIGK 63  
|||||  
DB 286 NNSGKDGNTSANSADSVKGPNTETNKKITDSNAVLAVKEVEALLSSIDEIAKAIGK 345  
OY 64 KIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGKLEKIDAKKCEFTTNK 123  
DB 346 KIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGKLEKIDAKKCEFTTNK 405  
OY 124 EKHTDLGKGVYDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 183  
DB 406 EKHTDLGKGVYDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 465  
OY 184 SPVVAESPKKPSMVNSGKDGNTSANSADSVK 216  
|||||  
DB 466 SPVVAESPKKPSMVNSGKDGNTSANSADSVK 498

RESULT 7  
US-08-235-836C-107  
Sequence 107, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-107

Query Match 49.5%; Score 915; DB 4; Length 466;  
Best Local Similarity 97.4%; Pred. No. 1.3e-61;  
Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

4 NNSGKDGNTSANSADSVKGPNTETNKKITDSNAVLAVKEVEALLSSIDEIAKAIGK 63  
|||||  
DB 276 NNSGKDGNTSANSADSVKGPNTETNKKITDSNAVLAVKEVEALLSSIDEIAKAIGK 335  
OY 64 KIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGKLEKIDAKKCEFTTNK 123  
DB 336 KIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGKLEKIDAKKCEFTTNK 395  
OY 124 EKHTDLGKGVYDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 183  
DB 396 EKHTDLGKGVYDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 455  
OY 184 SPVVAESPKK 194  
|||||  
DB 456 SPVVAESPKK 466

RESULT 8  
US-08-158-353-2  
Sequence 2, Application US/08158353  
Patent No. 5620662  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millida Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA

ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-2

Query Match 37.0%; Score 685; DB 1; Length 212;  
Best Local Similarity 74.1%; Pred. No. 1e-44;  
Matches 146; Conservative 19; Mismatches 28; Indels 4; Gaps 3;

QY 1 MACNSGKDGNTSANSADSVKPNLTETIKKTTDSNAVLAVKEVALLSSIDEIATA 60  
DB 17 ISCNNGSGGOSTANSDESVKGNLTETISKTTDSNAVLAVKEVETLVSSIDEIATA 76  
QY 61 ICKKHONNGDLENHNSGLAGAVASTLIKOKLDGLKN-EGKREIDAARCKSEFT 119  
DB 77 IGLILQ-NGLNAGANNGSLAGAVASTLIKOKLDGLKNSEELKREIDEAKCKNAFT 135  
QY 120 NKLKERTHDG--KEGYTDADAEALIKANGTKTGAEEELGKLFESVEVLSKAEMLAN 177  
DB 136 DRKSHAEELGANGATDANAKAILKTGTGDKGAELEKLFESVKNLSKAQETLNN 195  
QY 178 SVKELTSPVVAESPCKP 194  
DB 196 SVKELTSPVVAENPKRP 212

## RESULT 9

US-08-031-295-2  
Sequence 2, Application US/08031295  
Patent No. 5530103  
GENERAL INFORMATION:  
APPLICANT: LIVEY, Ian  
APPLICANT: DORNER, Friedrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
TITLE OF INVENTION: PREVENTION OF LYME DISEASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/031,295  
FILING DATE: 19930312  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/903,580  
FILING DATE: 25-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/142 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-031-295-2

Query Match 36.28%; Score 670; DB 1; Length 212;  
Best Local Similarity 73.0%; Pred. No. 1.4e-43;  
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSADSVKPNLTETIKKTTDSNAVLAVKEVALLSSIDEIATA 59  
DB 17 ISCNNGSGGOSTANSDESVKGNLTETISKTTDSNAVLAVKEVETLVSSIDEIATA 76  
QY 60 AICKKHONNGDLENHNSGLAGAVASTLIKOKLDGLKN-EGKREIDAARCKSEFT 118  
DB 77 AIGKRIQONNGLNAGNADKNSGLAGAVASTLIKOKLDGLKNSEELKREIDEAKCKSEFT 136  
QY 119 TNKLKERTHDGKEGYTDADAEALIKANGTKTGAEEELGKLFESVEVLSKAEMLAN 178  
DB 137 TKLIAAGHAQIGDGTNDSDKEALIKTGTGRTKGAEEELGKLFESVSLSKAQAESANS 196  
QY 179 VKELTSPVVAESPCKP 194  
DB 197 VKELTSPVVAETPKRP 212

## RESULT 10

US-07-903-580-2  
Sequence 2, Application US/07903580  
Patent No. 6221363  
GENERAL INFORMATION:  
APPLICANT: LIVEY, Ian  
APPLICANT: DORNER, Friedrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
TITLE OF INVENTION: PREVENTION OF LYME DISEASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/903,580  
FILING DATE: 19920625  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/131 IMM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-903-580-2

Query Match 36.2%; Score 670; DB 4; Length 212;  
Best Local Similarity 73.0%; Pred. No. 1.4e-43;  
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

QY 1 MACNNSGKDNLT-SANSADSVKGNLTETINKKITDTSNAVLAVKEVEALLSIDETIAK 59  
DB 17 ISCNNSGKGDSASTNPADSAKGNLTETISKKITDTSNAVLAVKEVEETLVSSIDELATK 76  
QY 60 AIGKKIHONNGDTEENNHNHSLAGAVAISTLIKOKIDGLKN-LEKIDAKKCSSEF 118  
DB 77 AIGKKIHONNGDTEENNHNHSLAGAVAISTLIKOKIDGLKN-LEKIDAKKCSSEF 136  
QY 119 TNKLKEXHTDLGKEGYTDADAKETIKANGTKTKGAEEGLKESEVYLSKAKEMLANS 178  
DB 137 TKTLAGHAQLGIDGATDNDNSKRAILKTNGTKTKGAEEGLKESEVYLSKAKEMLANS 196  
QY 179 VKELTSPVAESPKRP 194  
DB 197 VKELTSPVAETPKRP 212

RESULT 11  
US-08-158-353-4  
Sequence 4, Application US/08158353  
Patent No. 5620862

GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-4

Query Match 34.1%; Score 631; DB 1; Length 212;  
Best Local Similarity 69.9%; Pred. No. 1.3e-40;  
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNNSGKDNLT-SANSADSVKGNLTETINKKITDTSNAVLAVKEVEALLSIDETIAK 59  
DB 17 ISCNNSGKGDSASTNPADSAKGNLTETISKKITDTSNAVLAVKEVEETLVSSIDELATK 76  
QY 60 AIGKKIHONNGDTEENNHNHSLAGAVAISTLIKOKIDGLKN-LEKIDAKKCSSEF 118  
DB 77 AIGKKIHONNGDTEENNHNHSLAGAVAISTLIKOKIDGLKN-LEKIDAKKCSSEF 136  
QY 119 TNKLKEXHTDLGKEGYTDADAKETIKANGTKTKGAEEGLKESEVYLSKAKEMLANS 178  
DB 137 TKTLAGHAQLGIDGATDNDNSKRAILKTNGTKTKGAEEGLKESEVYLSKAKEMLANS 196  
QY 179 VKELTSPVAESPKRP 194  
DB 197 VKELTSPVAETPKRP 212

RESULT 12  
US-09-196-293-11  
Sequence 11, Application US/09196293  
Patent No. 6183755

GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Sautscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
TITLE OF INVENTION: burgdorferi  
FILE REFERENCE: 738.001052  
CURRENT APPLICATION NUMBER: US/09/196,293  
EARLIER FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 34.0%; Score 629; DB 4; Length 212;  
Best Local Similarity 70.4%; Pred. No. 1.8e-40;  
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

QY 1 MACNNSGKDG-NTSANSADSVKGNLTETINKKITDTSNAVLAVKEVEALLSIDETIAK 59

DB 17 ISCNNGKVGILSTPNPADESAGPVLTEISKITDSNMFVLAKEVEVLSTIDELAKK 76  
QY 60 AIGKRIHONNGLDTENNHNGSLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCKSETF 118  
DB 77 AIGKIDNNNGGLAALNNGNSLAGAVASTLITEKLSKLNLEELKTEIAKAKCKSEEF 136  
QY 119 TNKLEKHTDLGKGVTDADAKAELIKANGTKTKGAELKLFESVEVLSKAKEMLANS 178  
DB 137 TNKLSGHDLDGKODATDHAKAAILKTHATTDKGAKEFKDLFESVGLKKAQVALTNS 196  
QY 179 VKELTSPVAESPKKP 194  
DB 197 VKELTSPVAESPKKP 212

RESULT 13  
US-08-209-603E-11  
Sequence 11, Application US/08209603E  
Patent No. 6248538

GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILSK, BETTINA  
APPLICANT: PREAC-MURSEC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 357-5635

INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN  
DESCRIPTION: N/A  
HYPOTHETICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: B. BURGDORFERI  
IMMEDIATE SOURCE:  
LIBRARY: DSM 5662  
POSITION IN GENOME: N/A  
FEATURE:

IDENTIFICATION METHOD: amino acid analysis  
APPLICATION INFORMATION: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: INTERNAL  
ORIGINAL SOURCE:  
US-08-209-603E-11

Query Match 34.0%; Score 629; DB 4; Length 212;  
Best Local Similarity 70.4%; Pred. No. 1.8e-40;  
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

QY 1 MACNNGKDG-NTSANGSDSVKGPVLTEINKTKTDSNAVLLAVKEVEALLSIDELAKK 59  
DB 17 ISCNNGKVGILSTPNPADESAGPVLTEISKITDSNMFVLAKEVEVLSTIDELAKK 76  
QY 60 AIGKRIHONNGLDTENNHNGSLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCKSETF 118  
DB 77 AIGKIDNNNGGLAALNNGNSLAGAVASTLITEKLSKLNLEELKTEIAKAKCKSEEF 136  
QY 119 TNKLEKHTDLGKGVTDADAKAELIKANGTKTKGAELKLFESVEVLSKAKEMLANS 178  
DB 137 TNKLSGHDLDGKODATDHAKAAILKTHATTDKGAKEFKDLFESVGLKKAQVALTNS 196  
QY 179 VKELTSPVAESPKKP 194  
DB 197 VKELTSPVAESPKKP 212

RESULT 14  
US-08-235-836C-34  
Sequence 34, Application US/08235836C  
Patent No. 6248562

GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-235-836C-34

	Query Match Similarity	33.8%;	Score 626;	DB 4;	Length 212;
	Best Local Similarity	69.4%;	Pred. No. 3e-40;		
	Matches 136;	Conservative 16;	Mismatches 42;	Indels 2;	Gaps 2;
OY	1 MACNNSGKDGNT--SANSADSVKGPNLTIEINKITDINSNAVLAVKEVEALLTSIDEIANK	59			
Db	17 ISCSNGSGGDSASNPDESAKGCENLTFISKRIDSNAFVLAVERVELTVSIDELIAKK	76			
OY	60 AIGKRIHONNGGLDTENNHNHSGILLAGAYVISTILIKRKLDLGN--EGIKERIDAKKCESETE	118			
Db	77 AIGQVIDNNNGIALANNONGSLLAGAYVASTILTETLKSLKNLEELKTETIAAKKCSSEFF	136			
OY	119 TNKLIKERTLDLGKEGVTDADAKKALILKANGTTRKAEEIKGFESSEVEYLSSRAKKEMLNS	178			
Db	137 TNKLTSGHADJLCKODATDDHAFAAILTKTHAFTDKRAKKEFKOLFESVEEGLIRAAQVALTNS	196			
	179 VKELTSPVAASEPKP	194			
	197 VKELTSPVAASEPKP	212			

[illegible]

Search completed: March 18, 2002, 09:55:33  
Job time: 346 sec

```

RESULT 15
US-08-235-836C-32
; Sequence 32, Application US/08235836C
; Patent No. 6248562
;
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patencln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
TELECOMMUNICATION INFORMATION: BNL93-28A
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ. ID NO.: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-32

```

```
Query Match      33.2%   Score 614.5; DB 4; Length 209,  
Best Local Similarity    68.2%   Pred. No. 2..2e-39;  
Matches 133; Conservative 20; Mismatches 39; Indels 3; Gaps 2.  
  
QY 1 MACNNSGKGDTANSASAEVSGPMLTEINKTKITDSNVAVLAKREVELSSIDEELAAKA 60  
::||||| :|| :: || | ||||| : ||||| ||||| ||||| |||||
```



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:43 ; Search time 621.2 Seconds

(without alignments)  
168.507 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850  
Sequence: 1 MACNSGKDGNTSANSADSE.....AVENLAKRAKEMLANSVKEL 377

## Scoring table:

BIOSUM62  
Gap 10.0 , gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	377	19	US-09-596-746-30
2	1850	100.0	377	19	US-09-596-746A-30
3	1825	98.6	400	19	US-09-596-746-82
4	1825	98.6	401	19	US-09-596-746A-82
5	1801.5	97.4	374	19	US-09-596-746A-26
6	1796.5	97.1	373	19	US-09-596-746-26
7	1793.5	96.9	397	19	US-09-596-746-62
8	1793.5	96.9	398	19	US-09-596-746A-62
9	1614.5	87.3	378	19	US-09-596-746-28

10	1614.5	87.3	378	19	US-09-596-746A-28	Sequence 28, Appl
11	1589.5	85.9	401	19	US-09-596-746-80	Sequence 80, Appl
12	1589.5	85.9	402	19	US-09-596-746A-80	Sequence 80, Appl
13	1566	84.6	375	19	US-09-596-746A-24	Sequence 24, Appl
14	1561	84.4	374	19	US-09-596-746-24	Sequence 24, Appl
15	1558	84.2	398	19	US-09-596-746-60	Sequence 60, Appl
16	1558	84.2	399	19	US-09-596-746A-60	Sequence 60, Appl
17	1524.5	82.4	370	19	US-09-596-746A-36	Sequence 36, Appl
18	1519.5	82.1	369	19	US-09-596-746-36	Sequence 36, Appl
19	1516.5	82.0	368	19	US-09-596-746-66	Sequence 66, Appl
20	1516.5	82.0	394	19	US-09-596-746A-66	Sequence 66, Appl
21	1485.5	80.3	368	19	US-09-596-746A-42	Sequence 42, Appl
22	1480.5	80.0	367	19	US-09-596-746-42	Sequence 42, Appl
23	1477.5	79.9	391	19	US-09-596-746-74	Sequence 74, Appl
24	1477.5	79.9	392	19	US-09-596-746A-74	Sequence 74, Appl
25	1472	79.6	385	19	US-09-596-746A-56	Sequence 56, Appl
26	1467	79.3	384	19	US-09-596-746-56	Sequence 56, Appl
27	1465	79.2	560	19	US-09-596-746-52	Sequence 52, Appl
28	1465	79.2	560	19	US-09-596-746A-52	Sequence 52, Appl
29	1464	79.1	408	19	US-09-596-746-78	Sequence 78, Appl
30	1464	79.1	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289	69.7	371	19	US-09-596-746A-34	Sequence 34, Appl
32	1284	69.4	370	19	US-09-596-746-34	Sequence 34, Appl
33	1281	69.2	394	19	US-09-596-746-64	Sequence 64, Appl
34	1281	69.2	395	19	US-09-596-746A-64	Sequence 64, Appl
35	1264.5	68.4	370	19	US-09-596-746A-32	Sequence 32, Appl
36	1259.5	68.1	369	19	US-09-596-746-32	Sequence 32, Appl
37	1256.5	67.9	368	19	US-09-596-746-68	Sequence 68, Appl
38	1256.5	67.9	394	19	US-09-596-746A-68	Sequence 68, Appl
39	1250	67.6	369	19	US-09-596-746A-40	Sequence 40, Appl
40	1245	67.3	368	19	US-09-596-746-40	Sequence 40, Appl
41	1242	67.1	392	19	US-09-596-746-72	Sequence 72, Appl
42	1242	67.1	393	19	US-09-596-746A-72	Sequence 72, Appl
43	1225.5	66.2	368	19	US-09-596-746-38	Sequence 38, Appl
44	1225.5	66.2	369	19	US-09-596-746A-38	Sequence 38, Appl
45	1217.5	65.8	391	19	US-09-596-746-76	Sequence 76, Appl

## ALIGNMENTS

RESULT 1  
US-09-596-746-30  
; Sequence 30, Application US/09596746  
; GENERAL INFORMATION:  
; APPLICANT: Datwyler, Raymond J.  
; APPLICANT: Seinst, Gerald  
; APPLICANT: Dykhuisen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746-30

Query Match 100.0%; Score 1850; DB 19; Length 377;  
Best local similarity 100.0%; Pred. No. 3, 4e-125;  
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MACNSGKDGNTSANSADSEYKGNLTIKKTIDSNVLLAVEVALLSIDEIAKA 60

```

Db      1  MACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLLAVEEALLSIDEIAKA 60
Qy      61  ICKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Db      61  ICKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Qy      121  KLEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 180
Db      121  KLEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 180
Qy      181  ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 240
Db      181  ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 240
Qy      241  ETLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 300
Db      241  ETLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 300
Qy      301  EKTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 360
Db      301  EKTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 360
Qy      361  NLAKAEMLANSVKEL 377
Db      361  NLAKAEMLANSVKEL 377

```

## RESULT 2

```

US-09-596-746a-30
; Sequence 30, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30

```

```

Query Match      100.0%; Score 1850; DB 19; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.4e-125;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLLAVEEALLSIDEIAKA 60
Db      1  MACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLLAVEEALLSIDEIAKA 60
Qy      61  ICKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Db      61  ICKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Qy      121  KLEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 180
Db      121  KLEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 180
Qy      181  ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 240
Db      181  ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 240

```

```

Db      181  ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 240
Qy      241  ETLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 300
Db      241  ETLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 300
Qy      301  EKTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 360
Db      301  EKTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 360
Qy      361  NLAKAEMLANSVKEL 377
Db      361  NLAKAEMLANSVKEL 377

```

## RESULT 3

```

US-09-596-746-82
; Sequence 82, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 400
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-82

```

```

Query Match      98.6%; Score 1825; DB 19; Length 400;
Best Local Similarity 98.9%; Pred. No. 2.4e-123;
Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      2  ACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLLAVEEALLSIDEIAKA 61
Db      25  SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLLAVEEALLSIDEIAKA 84
Qy      62  GKKIHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 121
Db      85  GKKIHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 144
Qy      122  LKEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 181
Db      145  LKEKHTDLGKEGYTDADAKAELIKANGTKGAEEGLKFESEVYLSKAEMLANSVK 204
Qy      182  LSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 241
Db      205  LSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAKEI 264
Qy      242  TLLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 301
Db      265  TLLASIDELATRAIGKIKOONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSERIK 324
Qy      302  KTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 361
Db      325  KTENAKCSDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKFEKAVE 384
Qy      362  LAKAKEMLANSVKEL 377
Db      385  LAKAKEMLANSVKEL 400

```

```
RESULT 4
US-09-596-746a-82
; Sequence 82, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospc Chimera
-09-596-746a-82

Query Match          98.6%; Score 1825; DB 19; Length 401;
Best Local Similarity 98.9%; Pred. No. 2,4e-123;
Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGNLFEINKITDSNAVLLAVEVEALLSSIDEIAAKAI 61
DB 26 SCNNSKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEALLSSIDEIAAKAI 85
OY 62 GKRIHONGIDETENNNGSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTNK 121
DB 86 GKRIHONGIDETENNNGSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTNK 145
OY 122 LKEKHHDLDGEGVTDADAKFAILKANGTKTKGAEBELGKLFESVEVLSKAEMLANSVKE 181
DB 146 LKEKHHDLDGEGVTDADAKFAILKANGTKTKGAEBELGKLFESVEVLSKAEMLANSVKE 205
OY 182 LMSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEIE 241
DB 206 LMSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEIE 265
OY 242 TLLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSERLKE 301
DB 266 TLLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSERLKE 325
OY 302 KIENAKKSEDFTFKLEGEHAQGIENVTDENAKKAILITDAKKGAELKLEKFAVEN 361
DB 326 KIENAKKSEDFTFKLEGEHAQGIENVTDENAKKAILITDAKKGAELKLEKFAVEN 385
OY 362 LAKAAKEMLANSVKEL 377
DB 386 LAKAAKEMLANSVKEL 401

RESULT 5
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
US-09-596-746a-26

Query Match          97.4%; Score 1801.5; DB 19; Length 374;
Best Local Similarity 98.1%; Pred. No. 1.1e-121;
Matches 370; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSVKGNLFEINKITDSNAVLLAVEVEALLSSIDEIAKA 60
DB 1 MACNNSGKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEALLSSIDEIAKA 60
OY 61 GKRIHONGIDETENNNGSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTN 120
DB 61 GKRIHONGIDETENNNGSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTN 120
OY 121 KLEKHHDLDGEGVTDADAKFAILKANGTKTKGAEBELGKLFESVEVLSKAEMLANSVK 180
DB 121 KLEKHHDLDGEGVTDADAKFAILKANGTKTKGAEBELGKLFESVEVLSKAEMLANSVK 180
OY 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEI 240
DB 181 ELTSPVVAES---PAMVNNNSGKDGNTSANSADSVKGNLFEISKITDSNAVLLAVEVEI 237
OY 241 ETLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSERL 300
DB 238 ETLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSERL 297
OY 301 EKINAKKSEDFTFKLEGEHAQGIENVTDENAKKAILITDAKKGAELKLEKFAVE 360
DB 298 EKINAKKSEDFTFKLEGEHAQGIENVTDENAKKAILITDAKKGAELKLEKFAVE 357
OY 361 NLAKAAKEMLANSVKEL 377
DB 358 NLAKAAKEMLANSVKEL 374

RESULT 6
US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
US-09-596-746-26

Query Match          97.1%; Score 1796.5; DB 19; Length 373;
```

Best Local Similarity 98.1%; Pred. No. 2,4e-121;  
Matches 369; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 61
Db 1 ACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 60
OY 62 GKRIHONGLDTEENNHSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 121
Db 61 GKRIHONGLDTEYNNHNSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 120
OY 122 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
Db 121 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 180
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEKIE 241
Db 181 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEIE 237
242 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 301
Db 238 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 297
OY 302 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 361
Db 298 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 357
OY 362 LAKAKEMLANSVKEL 377
Db 358 LAKAKEMLANSVKEL 373
```

## RESULT 7

```
US-09-596-746-62
; Sequence 62, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746-62
```

Query Match 96.9%; Score 1793.5; DB 19; Length 397;  
Best Local Similarity 97.9%; Pred. No. 4,4e-121;  
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 84
OY 62 GKRIHONGLDTEENNHSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 121
Db 85 GKRIHONGLDTEYNNHNSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 144
OY 122 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
Db 145 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 204
```

```
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEKIE 241
Db 205 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEKIE 261
OY 242 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 301
Db 262 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 321
OY 302 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 361
Db 322 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 381
OY 362 LAKAKEMLANSVKEL 377
Db 382 LAKAKEMLANSVKEL 397
```

## RESULT 8

```
US-09-596-746a-62
; Sequence 62, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746a-62
```

Query Match 96.9%; Score 1793.5; DB 19; Length 398;  
Best Local Similarity 97.9%; Pred. No. 4,4e-121;  
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 61
Db 26 SCNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAVLLAVEKVEALLSSIDEIAAKAI 85
OY 62 GKRIHONGLDTEENNHSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 121
Db 86 GKRIHONGLDTEYNNHNSILAGAYASTLIKOKLDGKNEGKKEIDAAKCEFTFNK 145
OY 122 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
Db 146 LKEKHTDLDKEGYTDADAKAAILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 205
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEKIE 241
Db 206 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKKITESNAVLLAVEIE 262
242 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 301
Db 263 TLLASIDELATKAIGKKTIOONGGLAVEGHNGTLLAGAYTISKLTOKLDGKNSKEKLE 322
OY 302 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 361
Db 323 KIENAKKCEDEFTKKEGHAOLGIENTVDENAKKAILITDAADKGAELKFKAVEN 382
OY 362 LAKAKEMLANSVKEL 377
Db 383 LAKAKEMLANSVKEL 398
```

```
RESULT 9
US-09-596-746-28
; Sequence 28, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match      87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
QY 61 IGGKIHONNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
DB 61 IGGKIHONNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
QY 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
QY 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
DB 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
QY 241 ETLTASIDELATKAIGKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
DB 241 ETLTASIDELATKAIGKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
QY 301 EKIEENKCKSEDTTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALEKLFKAVE 360
DB 301 EKIEENKCKSEDTTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALEKLFKAVE 360
QY 360 NLKRAKAKEMLTNSVKEL 376
DB 360 NLKRAKAKEMLTNSVKEL 376

RESULT 10
US-09-596-746A-28
; Sequence 28, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
```

```
FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match      87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
QY 61 IGGKIHONNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
DB 61 IGGKIHONNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
QY 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
QY 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGYTDADAEALIKANGTKGAEELKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
DB 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
QY 241 ETLTASIDELATKAIGKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
DB 241 ETLTASIDELATKAIGKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
QY 301 EKIEENKCKSEDTTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALEKLFKAVE 360
DB 301 EKIEENKCKSEDTTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALEKLFKAVE 360
QY 360 NLKRAKAKEMLTNSVKEL 376
DB 360 NLKRAKAKEMLTNSVKEL 376

RESULT 11
US-09-596-746-80
; Sequence 80, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746-80
```

Query Match 85.9%; Score 1589.5; DB 19; Length 401;  
 Best Local Similarity 86.7%; Pred. No. 2.5e-106;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDNTSANSADSVKGNLTETIKKTTDSNAVLAVKEEALSSIDEITAARAI 61  
 DB 25 SCNNSGKDNTSANSADSVKGNLTETISKKTTDSNAVLAVKEEALSSIDEITAARAI 84  
 QY 62 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 121  
 DB 85 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 144  
 QY 122 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 181  
 DB 145 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 204  
 QY 182 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 241  
 DB 205 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 264  
 QY 242 TLASIDELATRAIKGKIQONGGLAVEAGHNGTLAGAYTISKLITOKLOGKSEKIKE 301  
 DB 265 TLTSIDELA-KAIGKRIKNDVSLDNEADHNSLISGAYLISNLTIKKISAKIDSGELKA 323  
 QY 302 KIENAKKCEDEFTKKEGEHQAOLGIEVNTDENAKKAILITDAKDKGALEKLEKRAVEN 361  
 DB 324 EIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTKGADELKLEKESVK 383  
 QY 362 LAKAAKEMLANSVKEL 377  
 DB 384 LSKAAKEMLTNSVKEL 399

## RESULT 12

US-09-596-746a-80  
 ; Sequence 80, Application US/09596746a  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dattwyler, Raymond J.  
 ; APPLICANT: Seimost, Gerald  
 ; APPLICANT: Dykhuizen, Daniaal  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746a  
 ; PRIOR FILING DATE: 2000-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 80  
 ; LENGTH: 402  
 ; TYPE: PRT  
 ; ORGANISM: ospc Chimera  
 ; US-09-596-746a-80

Query Match 85.9%; Score 1589.5; DB 19; Length 402;  
 Best Local Similarity 86.7%; Pred. No. 2.5e-106;  
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDNTSANSADSVKGNLTETIKKTTDSNAVLAVKEEALSSIDEITAARAI 61  
 DB 26 SCNNSGKDNTSANSADSVKGNLTETISKKTTDSNAVLAVKEEALSSIDEITAARAI 85  
 QY 62 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 121  
 DB 86 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 145  
 QY 122 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 181

DB 146 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 205  
 QY 182 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 241  
 DB 206 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 265  
 QY 242 TLASIDELATRAIKGKIQONGGLAVEAGHNGTLAGAYTISKLITOKLOGKSEKIKE 301  
 DB 266 TLTSIDELA-KAIGKRIKNDVSLDNEADHNSLISGAYLISNLTIKKISAKIDSGELKA 324  
 QY 302 KIENAKKCEDEFTKKEGEHQAOLGIEVNTDENAKKAILITDAKDKGALEKLEKRAVEN 361  
 DB 325 EIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTKGADELKLEKESVK 384  
 QY 362 LAKAAKEMLANSVKEL 377  
 DB 385 LSKAAKEMLTNSVKEL 400

## RESULT 13

US-09-596-746a-24  
 ; Sequence 24, Application US/09596746a  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dattwyler, Raymond J.  
 ; APPLICANT: Seimost, Gerald  
 ; APPLICANT: Dykhuizen, Daniaal  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746a  
 ; PRIOR FILING DATE: 2000-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 24  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Ospc Chimera  
 ; US-09-596-746a-24

Query Match 84.8%; Score 1566; DB 19; Length 375;  
 Best Local Similarity 85.9%; Pred. No. 1.1e-104;  
 Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDNTSANSADSVKGNLTETIKKTTDSNAVLAVKEEALSSIDEITAARAI 60  
 DB 1 MACNNSGKDNTSANSADSVKGNLTETISKKTTDSNAVLAVKEEALSSIDEITAARAI 60  
 QY 61 IGKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTN 120  
 DB 61 IGKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTN 120  
 QY 121 KLEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVK 180  
 DB 121 KLEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVK 180  
 QY 181 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 240  
 DB 181 LTPVVAESPPKPSMVNNNSGKDNTSANSADSVKGNLTETISKKTTESNAVLAVKEIE 240  
 QY 241 TLASIDELATRAIKGKIQONGGLAVEAGHNGTLAGAYTISKLITOKLOGKSEKIKE 300  
 DB 238 TLTSIDELA-KAIGKRIKNDVSLDNEADHNSLISGAYLISNLTIKKISAKIDSGELK 296  
 QY 301 KIENAKKCEDEFTKKEGEHQAOLGIEVNTDENAKKAILITDAKDKGALEKLEKRAVE 360  
 DB 297 EIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTKGADELKLEKESVK 356

OY 361 NLAKAKEMLANSVKEL 377  
 ||:|||||  
 Db 357 NLSKAKEMLTNSVKEL 373

RESULT 14  
 US-09-596-746-24

; Sequence 24, Application US/09596746  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Datwyler, Raymond J.  
 ; APPLICANT: Seinoost, Gerald  
 ; APPLICANT: Dykhuitzen, Dantale  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746  
 ; CURRENT FILING DATE: 2000-06-16  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; PRIOR FILING DATE: 1999-06-18  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 24  
 ; LENGTH: 374  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: OspC Chimera  
 US-09-596-746-24

Query Match 84.4%; Score 1561; DB 19; Length 374;  
 Best Local Similarity 85.9%; Pred. No. 2,6e-104;  
 Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVEALLSSIDEIAAKAI 61  
 ||:|||||  
 Db 1 ACNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVEALLSSIDEIAAKAI 60  
 OY 62 GKRIHONGIDTEENNNGSLAGAYASTLIKOKLDLKNEGLEKEIDAAKCCSEPTNK 121  
 ||:|||||  
 Db 61 GKRIHONGIDTEENNNGSLAGAYASTLIKOKLDLKNEGLEKEIDAAKCCSEPTNK 120  
 OY 122 LKKEHTDLGEGVTDADAKAAILKANGTKYGAELGKLFESVEVLSKAKEMLANSVKE 181  
 ||:|||||  
 Db 121 LKKEHTDLGEGVTDADAKAAILKANGTKYGAELGKLFESVEVLSKAKEMLANSVKE 180  
 OY 182 LTPVVAESPKPSMVNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVE 241  
 ||:|||||  
 Db 181 LTPVVAES---PAMVNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVE 237  
 OY 242 TLLASIDELATKAIGKIIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDLKNSEKLE 301  
 ||:|||||  
 Db 238 TLLTSIDELA-KAIGKIKIKDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKA 296  
 OY 302 KIENAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVEN 361  
 ||:|||||  
 Db 297 EIEKAKKCEDEFTAKLGEHTDLGEGVTDONAKKAILITNDKTKGADELEKLFESVK 356  
 OY 362 LAKAKEMLANSVKEL 377  
 ||:|||||  
 Db 357 LSKAKEMLTNSVKEL 372

RESULT 15  
 US-09-596-746-60  
 ; Sequence 60, Application US/09596746  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Datwyler, Raymond J.  
 ; APPLICANT: Seinoost, Gerald  
 ; APPLICANT: Dykhuitzen, Dantale

; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746  
 ; CURRENT FILING DATE: 2000-06-16  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; PRIOR FILING DATE: 1999-06-18  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 60  
 ; LENGTH: 398  
 ; TYPE: PRT  
 ; ORGANISM: OspC Chimera  
 US-09-596-746-60

Query Match 84.2%; Score 1558; DB 19; Length 398;  
 Best Local Similarity 85.6%; Pred. No. 4.7e-104;  
 Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVEALLSSIDEIAAKAI 61  
 ||:|||||  
 Db 25 ACNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVEALLSSIDEIAAKAI 84  
 OY 62 GKRIHONGIDTEENNNGSLAGAYASTLIKOKLDLKNEGLEKEIDAAKCCSEPTNK 121  
 ||:|||||  
 Db 85 GKRIHONGIDTEENNNGSLAGAYASTLIKOKLDLKNEGLEKEIDAAKCCSEPTNK 144  
 OY 122 LKKEHTDLGEGVTDADAKAAILKANGTKYGAELGKLFESVEVLSKAKEMLANSVKE 181  
 ||:|||||  
 Db 145 LKKEHTDLGEGVTDADAKAAILKANGTKYGAELGKLFESVEVLSKAKEMLANSVKE 204  
 OY 182 LTPVVAESPKPSMVNNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVE 241  
 ||:|||||  
 Db 205 LTPVVAES---PAMVNSGKDGNTSANSADSEYKGNLTETINKITDSNAVLAVKEVE 261  
 OY 242 TLLASIDELATKAIGKIIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDLKNSEKLE 301  
 ||:|||||  
 Db 262 TLLTSIDELA-KAIGKIKIKDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKA 320  
 OY 302 KIENAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVEN 361  
 ||:|||||  
 Db 321 EIEKAKKCEDEFTAKLGEHTDLGEGVTDONAKKAILITNDKTKGADELEKLFESVK 380  
 OY 362 LAKAKEMLANSVKEL 377  
 ||:|||||  
 Db 381 LSKAKEMLTNSVKEL 396

Search completed: March 18, 2002, 10:08:44  
 Job time: 972 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:19 ; Search time 81.02 Seconds  
(without alignments)  
297.254 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850

Sequence: 1 MACNNSGKDGTSANSADES.....AVENLAKAKEMLANSVKEL 377

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Search: 303921 seqs, 63882009 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents, AA, New.\*  
1: /cgn2\_6/ptodata/2/paa/BCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	692.5	37.4	211	6	US-09-974-992-5
2	631	34.1	212	6	US-09-974-992-7
3	596.5	32.2	207	6	US-09-974-992-3
4	172.5	9.3	1014	6	US-09-708-427-19883
5	172.5	9.3	1018	6	US-09-708-427-19882
6	172.5	9.3	1269	6	US-09-708-427-19881
7	168	9.1	630	6	US-09-742-096-5
8	160	8.6	1786	6	US-09-742-096-3
9	159.5	8.6	1639	6	US-09-614-150-6213
10	156.5	8.5	805	6	US-09-708-427-3730
11	156.5	8.5	841	6	US-09-708-427-3729
12	155.5	8.4	1038	6	US-09-708-427-32624
13	155.5	8.4	1042	6	US-09-708-427-32623
14	155.5	8.4	1043	6	US-09-708-427-32622
15	153	8.3	1881	7	US-10-032-585-7646
16	153	8.3	1881	7	US-10-032-585-7646
17	152.5	8.2	2437	6	US-09-815-242-5834
18	152.5	8.2	2437	6	US-09-815-242-5834
19	152.5	8.2	6281	6	US-09-815-242-12996
20	152.5	8.2	6281	6	US-09-815-242-12996
21	152	8.2	2086	6	US-09-815-242-5639
22	152	8.2	2086	6	US-09-815-242-5639
23	152	8.2	5795	6	US-09-815-242-12610
24	152	8.2	5795	6	US-09-815-242-12610
25	151.5	8.2	2056	6	US-09-614-150-4824

26	150.5	8.1	2434	6	US-09-815-242-5835	Sequence 5835, Ap
27	150.5	8.1	2434	7	US-10-072-851-5835	Sequence 5835, Ap
28	150	8.1	837	6	US-09-815-242-5883	Sequence 5883, Ap
29	150	8.1	837	7	US-10-072-851-5883	Sequence 5883, Ap
30	150	8.1	875	6	US-09-815-242-13080	Sequence 13080, A
31	150	8.1	875	7	US-10-072-851-13080	Sequence 13080, A
32	148.5	8.0	730	6	US-09-708-427-3731	Sequence 3731, Ap
33	148.5	8.0	1313	6	US-09-708-427-15044	Sequence 15044, A
34	148	8.0	3070	6	US-09-961-403-7	Sequence 7, Appl1
35	148	8.0	3110	6	US-09-561-7098-7	Sequence 7, Appl1
36	147.5	8.0	738	6	US-09-708-427-21267	Sequence 21267, A
37	147.5	8.0	763	6	US-09-708-427-21266	Sequence 21266, A
38	147.5	8.0	1304	6	US-09-708-427-15045	Sequence 15045, A
39	145.5	7.9	1690	6	US-09-614-150-10224	Sequence 10224, A
40	145.5	7.9	1690	6	US-09-614-150-10311	Sequence 10311, A
41	142.5	7.7	1144	6	US-09-708-427-15046	Sequence 15046, A
42	142	7.7	1489	6	US-09-614-150-6636	Sequence 6636, Ap
43	141.5	7.6	1725	7	US-10-037-182-12	Sequence 12, Appl
44	141.5	7.6	1786	5	US-09-938-275-7	Sequence 7, Appl1
45	141.5	7.6	1786	7	US-10-037-182-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-974-992-5  
: Sequence 5, Application US/09974992  
: GENERAL INFORMATION:  
: APPLICANT: Mathiesen, Marianne J.  
: APPLICANT: Thiesen, Michael  
: APPLICANT: Holm, Arne  
: APPLICANT: Ostergaard, Soren  
: TITLE OF INVENTION: Novel OspC-derived peptide fragments  
: FILE REFERENCE: 459-666P  
: CURRENT APPLICATION NUMBER: US/09/974,992  
: PRIOR APPLICATION NUMBER: 2001-10-10  
: PRIOR APPLICATION NUMBER: 09/180,089  
: PRIOR FILING DATE: 1999-05-13  
: PRIOR APPLICATION NUMBER: PCT/DK97/00203  
: PRIOR FILING DATE: 1997-05-02  
: NUMBER OF SEQ ID NOS: 40  
: SOFTWARE: Patent Ver. 2.1  
: SEQ ID NO 5  
: LENGTH: 211  
: TYPE: PRT  
: ORGANISM: Borrelia burgdorferi  
US-09-974-992-5

Query Match 37.4%, Score 692.5, DB 6, Length 211;  
Best Local Similarity 77.0%, Pred. No. 3.5e-43;  
Matches 151; Conservative 12; Mismatches 30; Indels 3; Gaps 3;

1 MACNNSGKDGTSANSADESVKGRPLTEINKRTIDSNVLLAVKEVALLSIDEIATAKA 60  
17 ISCNNSGKDGTSANSADESVKGRPLTEISKRTIDSNVLLAVKEVALLSIDEIATAKA 75  
61 IGRKTHONGIDTENNHNHSGSLAGVAISTLIKOKLDLK-NEGKLEKIDAKKSETFT 119  
76 IGRKTHONGIDTENNHNHSGSLAGVAISTLIKOKLDLK-NEGKLEKIDAKKSETFT 135  
120 NKLKEKHTDLCKEGVTDADAKKALIKANGT-KTGADELKGLFESVEVLSAKAKEMLAN 178  
136 TKLDNHNQGLIGVTDENNAKAKILKANAAGKDKGVELEKLSGSLSELSAKAKEMLAN 195  
179 VKELTSPVVAESPCKP 194  
196 VKELTSPVVAESPCKP 211  
RESULT 2  
US-09-974-992-7

```

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666p
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

```

```

Query Match          34.1%; Score 631; DB 6; Length 212;
Best Local Similarity 69.9%; Pred. No. 1e-38;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNNGSGKQNT-SANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSIDETIAAK 59
Db 17 ISCNNGSGKGDASTNPADSAKGNLTETSKRTIDSNAPLAVKEVEITLSDIELAKK 76
QY 60 AIGKTIHONGLDTENNHNSGLAGAVAIISLTIKOKLDGLN-EGLEKIDDAKCSETF 118
Db 77 AIGOKIDNNNGLAALNNGSLAGAVAIISLTIKLSKLNLELTETELAKKCSSEF 136
QY 119 TNKLEKHTDGLKEGVTADAKAELIKANGTKTGAEELGTFESVEVLSKAEMLANV 178
Db 137 TNKLSGHRDLKQDAPDHDHAKAILKTHATTDGAKFEKDFESVGLKAAQVALTNS 196
QY 179 VKELTSPVAESPKRP 194
Db 197 VKELTSPVAESPKRP 212

```

```

RESULT 3
US-09-974-992-3
; Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666p
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3

```

```

Query Match          32.2%; Score 596.5; DB 6; Length 207;
Best Local Similarity 66.0%; Pred. No. 3.2e-36;
Matches 128; Conservative 25; Mismatches 38; Indels 3; Gaps 2;

```

```

QY 1 MACNNGSGKQNT-SANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSIDETIAAK 60
Db 17 ISCNNGSGKGDASTNPADSAKGNLTETSKRTIDSNAPLAVKEVEITLSDIELAKK 76
QY 61 IGGKTIHONGLDTENNHNSGLAGAVAIISLTIKOKLDGLN-EGLEKIDDAKCSETF 120
Db 74 IGGKTIKNDGTLDNANNNESLIGAVIEISLTIKOKLSVINSSELEKIKIEADCSSEFTT 133
QY 121 KLEKHTDGLKEGVTADAKAELIKANGTKTGAEELGTFESVEVLSKAEMLANV 180
Db 134 KLEKSHAEELGTSYQVODNAAKAILKTHGTDKAKELBELFKSLSLSKAAQVALTNSVK 193
QY 181 ELTSPVAESPKRP 194
Db 194 ELTSPVAESPKRP 207

```

```

RESULT 4
US-09-708-427-19883
; Sequence 19883, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc-feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

```

```

Query Match          9.3%; Score 172.5; DB 6; Length 1014;
Best Local Similarity 23.2%; Pred. No. 0.00018;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KQNTSANSADSVKGPMLTEINKRTDSNAVLAVKEVEALLSIDETIAAK 45
Db 401 KHEGTADSKGYLGVAELQSTLEAFQVKSLSLEALNATENEKELEENAVTSEKK 460
QY 46 VEALLSIDETIAAK-----AIGKTIHONGLDTENNHNSGLAGAVAIISLTIK 94
Db 461 LEA---TVDEIVYKISESENNLESTIRNELNVTOG-KLESTINDKAGGLDESEVMEKLS 516
QY 95 KLDGLNBSLKEKIDAA-----KKCSSEFTNK-----LKEKH 126
Db 517 ABESELEQKG--REIDEATTKRMELEALHQSLSIDSEHRLQANAEFTSRSEASSLLEKI 574
QY 127 TDU-KHEGVTADAKAELIKANGTKTGAEELGTFESVEVLSKAEMLANVSKEL 182
Db 575 RDLBGRKTSYEQLEAASGKSSLEKLEQTLGRLAAESVNEKLMQEPDQAOEKLSQSS 634
QY 183 -TSPVAESPKRP-----PSMVNNGSGKQNTSANSADSVKGPMLTETSKKTIENAV 233
Db 635 SESELEAETNNOQIKTIQLEGLIGSGSVKEKTAIRLEBALE-----RPNKETESSDL 689
QY 234 VLAVKEITLISLDELATATG-----KTIQONG-----GLAVEA 269
Db 690 VEKKTHTENOIEEYKRLAHASGVADTRKVELEDALSKLNLESTIEELGANCQGLERKS 749
QY 270 GHNGTL-----LAGAVITSKLITOKLDGLK-NSKLEKELEAKKCSSEFTKRL-EG 319

```

Db 750 GDIAEYVNLKLNLELANHSGSEANLEQTKLSALAEKEDQTANLEASKTIEDTLTKQLTSEG 809  
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369  
Db 810 EKIQSOIEKLRVAAAEKSVL-----ESHPELEKTLSEVKAQDKENVENAATAASVAVAE 863  
QY 370 LANSYKE 376  
Db 864 LITSKLOE 870

## RESULT 5

US-09-708-427-19882  
; Sequence 19882, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19882  
; LENGTH: 1018  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1018  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1018  
; OTHER INFORMATION: Ceres Seq. ID 1836846  
US-09-708-427-19882

Query Match 9.3%; Score 172.5; DB 6; Length 1018;  
Best Local Similarity 23.2%; Pred. No. 0.00018;  
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KDCNTSANSAD-----SVKGPNL-----TEINKKITDS-NAVLLAVKE 45  
Db 405 KHGETEADSKGYIGQVAELOSTLEAFQVXSSSLAALNITATENEKELTENLNAVTSSEKK 464  
QY 46 VEAALLSIDEIAK-----AIGKIHONNGLDPTENNHNGLAGAVAISTL--IKQ 94  
Db 465 LEA---TVDEYSKISESENLESLINELNVTQG-KLESTENDLKAAGLOESEVMEKLS 520  
QY 95 KLDGLKNEGLKEKIDAA-----KKCSEFTTNK-----LKEKH 126  
Db 521 AESLSLEQKG--REIDEATTKRMELALHOSLIDSEHRLQKAMEEFTSRDSEASSLTEKL 578  
QY 127 TDL-GKEGYTDADAKKAILKANGTKTKGAELGKLFESVYVLSKAEMLANSVKEL-- 182  
Db 579 RDEGKIKSYEEOGLAEASGSSSLKEKLEOTLGRLLAAESVNEKLEQEPDOAEKSLQSS 638  
QY 183 -TSPVVAESPCK-----PSMVNNSGKDGNTSANSADSVKGPNLTEISKITTESNAV 233  
Db 639 SESELLAETNNOLKIKITQIEGLEIGSGSVKEKTRAKRLLEAIE-----RNNOKETESSDL 633  
QY 234 VLAVKEITLLASIDELATKAIG-----KKIQONG---GLAVEA 269  
Db 694 VEKLKTHENQIEEYKKLAAHSGVADTRKVELLEDALSKLKNLESTIEELGAKCOGLEKES 753  
QY 270 GHNGTL-----LAGAYITSKLITQKLDGLK-NSEKLIKETINAKKCSDFTKKL--EG 319  
Db 754 GDIAEYVNLKLNLELANHSGSEANLEQTKLSALAEKEDQTANLEASKTIEDTLTKQLTSEG 813  
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369  
Db 814 EKIQSOIEKLRVAAAEKSVL-----ESHPELEKTLSEVKAQDKENVENAATAASVAVAE 867

QY 370 LANSYKE 376  
Db 868 LITSKLOE 874

## RESULT 6

US-09-708-427-19881  
; Sequence 19881, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19881  
; LENGTH: 1269  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1269  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1269  
; OTHER INFORMATION: Ceres Seq. ID 1836845  
US-09-708-427-19881

Query Match 9.3%; Score 172.5; DB 6; Length 1269;  
Best Local Similarity 23.2%; Pred. No. 0.00024;  
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KDCNTSANSAD-----SVKGPNL-----TEINKKITDS-NAVLLAVKE 45  
Db 656 KHGETEADSKGYIGQVAELOSTLEAFQVXSSSLAALNITATENEKELTENLNAVTSSEKK 715  
QY 46 VEAALLSIDEIAK-----AIGKIHONNGLDPTENNHNGLAGAVAISTL--IKQ 94  
Db 716 LEA---TVDEYSKISESENLESLINELNVTQG-KLESTENDLKAAGLOESEVMEKLS 771  
QY 95 KLDGLKNEGLKEKIDAA-----KKCSEFTTNK-----LKEKH 126  
Db 772 AESLSLEQKG--REIDEATTKRMELALHOSLIDSEHRLQKAMEEFTSRDSEASSLTEKL 829  
QY 127 TDL-GKEGYTDADAKKAILKANGTKTKGAELGKLFESVYVLSKAEMLANSVKEL-- 182  
Db 830 RDEGKIKSYEEOGLAEASGSSSLKEKLEOTLGRLLAAESVNEKLEQEPDOAEKSLQSS 889  
QY 183 -TSPVVAESPCK-----PSMVNNSGKDGNTSANSADSVKGPNLTEISKITTESNAV 233  
Db 890 SESELLAETNNOLKIKITQIEGLEIGSGSVKEKTRAKRLLEAIE-----RNNOKETESSDL 944  
QY 234 VLAVKEITLLASIDELATKAIG-----KKIQONG---GLAVEA 269  
Db 945 VEKLKTHENQIEEYKKLAAHSGVADTRKVELLEDALSKLKNLESTIEELGAKCOGLEKES 1004  
QY 270 GHNGTL-----LAGAYITSKLITQKLDGLK-NSEKLIKETINAKKCSDFTKKL--EG 319  
Db 1005 GDIAEYVNLKLNLELANHSGSEANLEQTKLSALAEKEDQTANLEASKTIEDTLTKQLTSEG 1064  
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369  
Db 1065 EKIQSOIEKLRVAAAEKSVL-----ESHPELEKTLSEVKAQDKENVENAATAASVAVAE 1118  
QY 370 LANSYKE 376  
Db 1119 LITSKLOE 1125

RESULT 7

```

US-09-742-096-5
; Sequence 5, Application US/09742096
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007/3050DIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-5

```

```

Query Match          9.1%; Score 168; DB 6; Length 630;
Best Local Similarity 18.9%; Pred. No. 0.0002;
Matches 82; Conservative 104; Mismatches 144; Indels 104; Gaps 20;

```

```

QY 9 DGNTS-----ANSADSVK-----GNLITEINKKIDSNAVLAAVEEALLSID 54
DB 135 DGSAASSVEERSIASVDESIDSIEENVAPTEVEI---VAPTEVEIYAPSVESVAP 191
QY 55 EIAKAIGKIKHONGNDTENNINGSILAGAVASTLIKOKLDGLKNEGKEIDAKK 114
DB 192 ESVAAPSVESVAE-----VEESVAENVEIYAPSVESV- AENV 230
QY 115 SEFTNKKEKHTDGEVTDADAKAAILKANGTKGAEEL- GKLEESV-----EVL 167
DB 231 EESVAENVEESVAENVEESVAE- NVEESV-----AENVEIYAPTEVESVAPTEVEI 282
QY 168 SKAKKEMANSVKELTSPVAESPKRPSVANNNGKDGNTS-ANSADSVKGNLITE-ISK 225
DB 283 APVESVAETVETVEIYAPSVES- VAPSVESVAENVEESVAENVEESV- AENVEESVAE 340
QY 226 KITESNAVLAAVEIEETLTA-SIDELATKAIGKIKOONGGLAVEGHNGTLLAGAYT--I 282
DB 341 NVEESVA-----ENVEIYAPSVETVEIYAPTEVESVAENVAATMLSDNLSNLLGIEETE 395
QY 283 SKLITOKLDGLKNS-----EKLK-----EKINAKKSEDFTKLEGEHA 322
DB 396 KOSILNEIEVEKENVYTTILEKEVETTAESVTFNSIIEIEQENTITNDTIEKLELHE 455
QY 323 QI-----GIENVTDENAKKAL-----ITDAKDKCAAELEKLFKAVENLA 363
DB 456 NVLSALETOTSEEEKKEVIVIEVEKEEVATTLTIEVQAESEESTTEIFEENLEBNA 515
QY 364 KAKKEMANSVKEL 377
DB 516 VESNEKVAENLEKL 529

```

```

RESULT 8
US-09-742-096-3
; Sequence 3, Application US/09742096
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007/3050DIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06

```

```

; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

```

```

Query Match          8.6%; Score 160; DB 6; Length 1786;
Best Local Similarity 18.7%; Pred. No. 0.0031;
Matches 76; Conservative 97; Mismatches 167; Indels 66; Gaps 16;

```

```

QY 19 ESKGNLITEINKKIDSNAVLAAVEEALLSIDELAKAIGKIKHONGNDTENNHN 78
DB 588 EELVAPTEVEI---VAPTEVEIYAPSVESVAPSVESVEEENVEESVAENVEESVAENVE 644
QY 79 GSLAG-----AYAISTLIKOKLDGLKNEGKEIDAKKSEFTNKLKREKHTDGLK 133
DB 645 ESVAENVEIYAPTEVEIYAPTEVEIYAPSVESV- APSVEESVEEENVEESVAENVEES 702
QY 134 VTDADAKAAILKANGTKGAEELGKLEESVEVLSKAKKEMANSVKELTSPVAES--P 191
DB 703 VAE- NVEESV- AENVEESVAENV-----BEIYAPTEVEIYAPTEVEIYAPSVESVAP 753
QY 192 KPSVANNNGKDGNTSANSADSVKGNLITE-ISKKITESNAVLAAVEI-----ETIL 244
DB 754 SVEESVEEENVEE--SVANVEESV- AENVEESVAENVEES- VAPTEVEIYAPSVESVA 808
QY 245 ASIDELATKAIGKIKOON-----GLAVEGHNGTLLAGAYTISKLITOKLDGLKNS-- 296
DB 809 PSVESVAENVAATMLSDNLSNLLGIEETEIKDSIINEIEVEKENVYTTILEVEETTA 868
QY 297 -----EKLKEINAKKSEDFTKLEGEHAOL---GIENVTDENAKKAL----- 339
DB 869 ESVTFNSIIEIEQENTITNDTIEKLELHEENVLSALETOTSEEEKKEVIVIEVEKE 928
QY 340 -----ITDAKDKCAAELEKLFKAVENLA KAKKEMANSVKEL 377
DB 929 EVATTLTIEVQAEKSNATITEIFEENLEBNAVESNENVAENLEKL 974

```

```

RESULT 9
US-09-614-150-6213
; Sequence 6213, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637

```

PRIOR FILING DATE: 2000-03-23  
 NUMBER OF SEQ ID NOS: 43008  
 SOFTWARE: FASTSEQ for Windows version 4.0  
 SEQ ID NO: 6213  
 LENGTH: 1639  
 TYPE: PRP  
 ORGANISM: DROSOPHILA  
 US-09-614-150-6213

Query Match 8.6%; Score 159.5; DB 6; Length 1639;  
 Best Local Similarity 23.9%; Pred. No. 0.003;  
 Matches 105; Conservative 58; Mismatches 162; Indels 115; Gaps 19;

QY 2 ACNSGKDCNTSANSAD-----ESYKPNLIEINKITDSNAVLAKEVEALLSIDE 55  
 DB 1107 ARNSGDCGQTAEVYIDDLHKHLDSVR-EHLVSADKFOADANGEIDRARNQNTYILDQITE 1165  
 QY 56 IAKAIGKKIHQNGDTENNHNSLLAGAVASTLIKQKLDGLKNEGLKEKIDAAKKS 115  
 DB 1166 NAKKEL-----QALDLNDEGQALARKKESVEFGQ-----SEQISDISREA 1210  
 DB 116 EFTFNKLEK-EKHTDLGEGVTDADAKKALKAN-----GFKTKG--AEELGK 159  
 DB 1211 RALADKLESEAFDLKNA-----KDAKDAVEKAHQALAKSAIDQLKIGTELSEVGLSH 1266  
 QY 160 LRESVEVLSKAKE-----MLANSYKELTSPVVAESPCKPSKVNNSKDGNT 206  
 DB 1267 VKSOLCTVQTSKEALRKANEVYDFTALTLLNDVNROTOPEIDISQK-----KDAVA 1318  
 QY 207 SANSADSVKPNLIEISKITTESNAVLAKEIETLLASIDELATKAIKKIQ----- 261  
 DB 1319 ANERADELK--QITELS-----NSGELFPADEFTBOEL-----TEALLKRAEQOQLED 1365  
 QY 262 -----NGLAVAGHNGTLGAGVYTSKLTITQKLDGLKNSK-----KL 299  
 DB 1366 IELLERAKAHADKATKAVGODN-TLKEANNTEYKLAGFQSDVQRSESEKALQYTPNI 1424  
 QY 300 KEKIEKAKCSEDFTKKLEGEHAQDGIENVTDENAKKAIL--ITDAKDKGALEKLFK 357  
 DB 1425 EKEIQVAESLISQAEALDGANKN--ANEAKKNAQEAQKVAEQASKD-----AELLR 1475  
 QY 358 AVENLAKAKKEMLANSVKEL 377  
 DB 1476 RKANETKVAARNLREBADDL 1495

RESULT 10  
 US-09-708-427-3730  
 Sequence 3730, Application US/09708427

APPLICANT: N. ALEXANDROV et al.  
 TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
 FILE REFERENCE: 2750-1243P  
 CURRENT APPLICATION NUMBER: US/09/708,427  
 CURRENT FILING DATE: 2000-11-09  
 NUMBER OF SEQ ID NOS: 85364  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 3730  
 LENGTH: 805  
 TYPE: PRP  
 ORGANISM: Arabidopsis thaliana  
 NAME/KEY: misc\_feature  
 LOCATION: 1..805  
 OTHER INFORMATION: Xaa is any amino acid  
 NAME/KEY: misc\_feature  
 LOCATION: 1..805  
 OTHER INFORMATION: Ceres Seq. ID 1810877  
 US-09-708-427-3730

Query Match 8.5%; Score 156.5; DB 6; Length 805;  
 Best Local Similarity 20.4%; Pred. No. 0.0019;  
 Matches 86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

QY 21 VKGPNLIEINK-----KITDSNAVLAKE--VEAL-----LSSIDEIAAKIGKKIH 66  
 DB 50 VKHKELEVEKAEADGLGLENSRKMIIEEDRIRISALAEKLEELQKOSASLEEKLK 109  
 QY 67 QNNGDTENNHNSLLAGAVASTLIKQKLDGLKNEGLKEK-----DAKKS 115  
 DB 110 IS--DERYSTKTLALLSQALSONSVLEOKIKSL--BELSEKVSSELKALIVAEEGKKS 164  
 QY 116 --EFTFNKLEKHTDGRGVTADDAKE--AIIKA-----NGKTKGAEELKLF 162  
 DB 165 IOMOEYOEKSKLESSLNOSANSELEEDRLIALQGAHEDIGVSTRSVLEGLQF 224  
 QY 163 SVEVLKAAKEMT-----ANSYKELTSPVVAESPCKPS-----WVNSGKDGNT 206  
 DB 225 TSQKLEKAEKLEKDLAIGYKNSLEATLSVAMEKERDLESENINAMKELKSEERLEK 284  
 QY 207 SANSADSVK-----GPNLIEISKITTESNAVLAKEIETLLASIDEL 250  
 DB 285 QAREIDEATRTSIELEALHKHSELKYQKTMEDESSRDTFAKSILTEKSKDEEKIRVEYEGK 344  
 QY 251 ATKAIKKIQONGGLAVEAGHNGTLGAGVYTSKLTITQKLDGLKNSK--LKEKIEKAK 308  
 DB 345 LAECGQSISLQELDQSSAEN-ELLADTNNQIKIKIQLEGYDSEKETAIEKLNOKT 403  
 QY 309 CSDEFKTKLEGEHAQDGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAKE 368  
 DB 404 EANDLITKLSH-----ENVIEBH-KROYLEASGVADTKRVEVEALLLNTLESTIEE 456  
 QY 369 M 369  
 DB 457 L 457

RESULT 11  
 US-09-708-427-3729  
 Sequence 3729, Application US/09708427

APPLICANT: N. ALEXANDROV et al.  
 TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
 FILE REFERENCE: 2750-1243P  
 CURRENT APPLICATION NUMBER: US/09/708,427  
 CURRENT FILING DATE: 2000-11-09  
 NUMBER OF SEQ ID NOS: 85364  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 3729  
 LENGTH: 841  
 TYPE: PRP  
 ORGANISM: Arabidopsis thaliana  
 NAME/KEY: misc\_feature  
 LOCATION: 1..841  
 OTHER INFORMATION: Xaa is any amino acid  
 NAME/KEY: misc\_feature  
 LOCATION: 1..841  
 OTHER INFORMATION: Ceres Seq. ID 1810876  
 US-09-708-427-3729

Query Match 8.5%; Score 156.5; DB 6; Length 841;  
 Best Local Similarity 20.4%; Pred. No. 0.002;  
 Matches 86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

QY 21 VKGPNLIEINK-----KITDSNAVLAKE--VEAL-----LSSIDEIAAKIGKKIH 66  
 DB 86 VKHKELEVEKAEADGLGLENSRKMIIEEDRIRISALAEKLEELQKOSASLEEKLK 145  
 QY 67 QNNGDTENNHNSLLAGAVASTLIKQKLDGLKNEGLKEK-----DAKKS 115

Db 146 IS---DERYKTDALLSOLASQNSVLEQKLS--EELSEKVELSKALIYAEBRGKSS 200  
QY 116 ---ETFTNKLKERTHDLEKGEVTDADAKE---ALKA-----NGRTKGAEBELKLEF 162  
Db 201 IOMQYOEVSLESLSSWSSARNSLEEDLRIALQGAHEHDIGNVSTKRSVLEGLQIFQ 260  
QY 163 SVEVLSKAKEML-----ANSVKELTSPVAESPSPKPS-----AVNNSGKQONT 206  
Db 261 TSQLEKAEKELKDLAIOYKNSLEATLSVAMEKEHEDSENINAVAMEKKSSEERLEK 320  
QY 207 SANSADSEVK-----GPNLTISKITTESNAVAVLAKETITLASIDEL 250  
Db 321 QARELDEFTTTSIELEHAKHSEKLVOKTMEFSSRDTFAKSTKESKIDLEKIRYVEGK 380  
QY 251 ATKAIGKKIOONGILAWEAGHNGTLLAGAVTISKITQOKLDGLKNSBK--LKEIKENAK 308  
Db 381 LAECGGGLSLQEBELDOSSAEN--ELADDTNNQIKIKIQLEGIYDSKEFAIEKLNOKDT 439  
QY 309 CSEDTKLEGEHMOGLIENVTDENAKKAILITDARKGAAEELKLFKAVENLAKAKE 368  
Db 440 EAKDLITRLKSH-----ENVIEEH--KROYLEASGVADTRKVEVEALLKLTLESTIEE 492  
QY 369 M 369  
Db 493 L 493

RESULT 12  
US-09-708-427-32624  
; Sequence 32624, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32624  
; LENGTH: 1038  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1038  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1038  
; OTHER INFORMATION: Ceres Seq. ID 1834662  
US-09-708-427-32624

Query Match 8.4%; Score 155.5; DB 6; Length 1038;  
Best Local Similarity 25.0%; Pred. No. 0.0032;  
Matches 104; Conservative 53; Mismatches 166; Indels 73; Gaps 16;  
QY 7 GKDQNTSAN---SADESXKGPNTLEINKKITDSNAVLAKVEEALLSSIDEIAAKAIGK 63  
Db 372 GKEGNEFCNVISDAEKSQSVINIPVANHLOEPRN---IPVKLSEHLPKPEPT-----K 423  
QY 64 KIHONNGIDTENNNHNGSLAGA-----YAISTLIKOKLDGLKNEGIEKIDAK--KC 114  
Db 424 RIAKNEPVKSTKKEQSSSSSEASKIPVCLRVDPLEPRNGSKSVSHPRMKSKEETKI 483  
QY 115 SETFTNKLKERTHDLEKGEVTDADAKEAILKANGTKTKGAEBELKLFESYEVLSKAKEM 174  
Db 484 AAPSSKKAESRTVPEACNVKCEDANEMKMAEGSLNALRTEKG-----SVESNSNLQES 539  
QY 175 LANSVKELTSPVAESPSPKPSMVNNSGKQNTSANSADSEVKGPNTL--EISKITESNA 232  
Db 540 NGELIRKPEAKENREQPAKKSFTEE-----AARIIQSMYKRDYVRWPEPTKKLEIAT 593

QY 233 VLAIV-----KEITELASIDELATAKIGKKIOONGL-----AVEAGHNG--TLIA 277  
Db 594 VREOMGDVKKRIRALEASTQHIIE--KEIYVNGSLVNNLLKIDAVGLHPSTREFPK 650  
QY 278 GATTISKLITQKIDGLKNS-----EKLKEIKENAKCSEDTFKLGEHMOGLIEN--V 329  
Db 651 ALATELSSIODKIDSLKNSCASAKEKVAKEVOEIKSQPSD---SPVNLHSOLTEENKMY 707  
QY 330 TDENAKKAILI-----TDAKDKGAAEELKLFKAVENLAKAKAKEMLANS 373  
Db 708 SDTNLEKVLRLSPDEHPMSVNLRTDEKQASAEETEGIGLEFETLATDSKQATENA 763

RESULT 13  
US-09-708-427-32623  
; Sequence 32623, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32623  
; LENGTH: 1042  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..1042  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..1042  
; OTHER INFORMATION: Ceres Seq. ID 1834661  
US-09-708-427-32623

Query Match 8.4%; Score 155.5; DB 6; Length 1042;  
Best Local Similarity 25.0%; Pred. No. 0.0032;  
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;  
QY 7 GKDQNTSAN---SADESXKGPNTLEINKKITDSNAVLAKVEEALLSSIDEIAAKAIGK 63  
Db 376 GKEGNEFCNVISDAEKSQSVINIPVANHLOEPRN---IPVKLSEHLPKPEPT-----K 427  
QY 64 KIHONNGIDTENNNHNGSLAGA-----YAISTLIKOKLDGLKNEGIEKIDAK--KC 114  
Db 428 RIAKNEPVKSTKKEQSSSSSEASKIPVCLRVDPLEPRNGSKSVSHPRMKSKEETKI 487  
QY 115 SETFTNKLKERTHDLEKGEVTDADAKEAILKANGTKTKGAEBELKLFESYEVLSKAKEM 174  
Db 488 AAPSSKKAESRTVPEACNVKCEDANEMKMAEGSLNALRTEKG-----SVESNSNLQES 543  
QY 175 LANSVKELTSPVAESPSPKPSMVNNSGKQNTSANSADSEVKGPNTL--EISKITESNA 232  
Db 544 NGELIRKPEAKENREQPAKKSFTEE-----AARIIQSMYKRDYVRWPEPTKKLEIAT 597  
QY 233 VLAIV-----KEITELASIDELATAKIGKKIOONGL-----AVEAGHNG--TLIA 277  
Db 598 VREOMGDVKKRIRALEASTQHIIE--KEIYVNGSLVNNLLKIDAVGLHPSTREFPK 654  
QY 278 GATTISKLITQKIDGLKNS-----EKLKEIKENAKCSEDTFKLGEHMOGLIEN--V 329  
Db 655 ALATELSSIODKIDSLKNSCASAKEKVAKEVOEIKSQPSD---SPVNLHSOLTEENKMY 711  
QY 330 TDENAKKAILI-----TDAKDKGAAEELKLFKAVENLAKAKAKEMLANS 373  
Db 712 SDTNLEKVLRLSPDEHPMSVNLRTDEKQASAEETEGYGLFETLATDSKQATENA 767  
RESULT 14

```
US-09-708-427-32622
; Sequence 32622, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243p
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 32622
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Ceres Seq. ID 1834660
; -09-708-427-32622

Query Match      8.4%; Score 155.5; DB 6; Length 1043;
Best Local Similarity 25.0%; Pred. No. 0.0032;
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;

QY 7 GKDQNTSAN---SADESVGPNLTETINKKITDSDNAVLAVKEVEALLSIDETIAAIGK 63
DB 377 GKEGNECVLSDAEKSSVINIPVANHLOEPNN---IPYKLESHLPRTEPT-----K 428
QY 64 KIHQNGLDTENNHNSLLGA-----YAISTLIKOKLDGLKNGLEKIDAK--KC 114
DB 429 RIAKNPVASTKKEQSSSSSEASKLPVCLRVDPPLPKRNGSKSVSHPRMEKSETKI 488
QY 115 SEFTTKKLEKHTDLCGEGYTDADAKAELIKANGTKTGAEELGKLPESVEVLSKAKEK 174
DB 489 AAPLSKKKESRTVPACVNCDCEDANAEMKMGSLNALRTEKG---SVESNSNLQEESS 544
QY 175 LANSVVELSPVVAESPKKPSMVNNSGKDQNTSANSADSVKGNLT--EISKITESNA 232
DB 545 NGEIIPCEAKKEKREOPAKKSTFEF-----AARIIQSMYRGYDVRMEPIKKLEIAT 598
QY 233 VYLA---KEIETLASIDELATKAIGKKIQONGGL-----AVEAGHNG--TLA 277
DB 599 VREQMGDVKKRIETALASTDQHIE--KEIYVNGELVNNLLKLDAVEGLHPSIREFRK 655
QY 278 GAYTISKLTOKLDGLKNS-----EKLKEKIEKAKKCEDFTKKLEGEHAQIGEN--V 329
DB 656 ALATELSSYIDKLDLSLKNCSAESAKEAVEVEIKSOPSD--SPVNLHESQLTEENKAV 712
QY 330 TDENAKKAILI-----TDAKDKGAELKELFAVENLAKAKEMLAN 373
DB 713 SDNLEKVLRLSPPEHPMSVNLNTDEKQASAEETEGVGLFETLATDSQKATENA 768

RESULT 15
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jlang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
```

```
LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: 1881..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

Query Match      8.3%; Score 153; DB 7; Length 1881;
Best Local Similarity 23.8%; Pred. No. 0.011;
Matches 91; Conservative 59; Mismatches 146; Indels 86; Gaps 16;

QY 27 TEINKKITDSNAVLAVKEVEALLSIDETIAAIGKIH--QNGLDTENNHNSLLAG 84
DB 1529 TELKDRISE-----VEKERAMISENSEVIVIKESDIKLSIKINSKESKE--- 1577
QY 85 AYAISTLIKOKLDGLKNE--GLKEKIDAAKCCSETFTNKLKEKHTDLCRGVTDADAKKA 142
DB 1578 ---TTHNEOK-TSLKODIAKLSQDHESAOPTOLEKENOLKELKASLEKHNTESATSTE- 1632
QY 143 ILKANGTKTGAEELGKLPESVEVLSKAKEMLANSVKELTSPVVAESPKKPSMVNNSGK 202
DB 1633 -----EKNQIKELSETISLTKTELKTSGDALKQSOKEV-----KTLTKNSDT 1676
QY 203 DG-----MNSANSADSVKGNLTETISKITESNAVLAVKEIETL---LAS 246
DB 1677 ESKLEKOLELEKVKSDLOTADEKLKGTREIETALK-----SELETFVKNSGLSLT 1725
QY 247 IDELA--TKAIGKKIQONGGLAVEAGHNGTLGAVYISKLTOKLDGLKN--SEKLEK 302
DB 1726 TSELAALTFTVKSLEKEKEELOPLSGNKSKELEDYIQKHSIDISEKLALYDELKEKTKOF 1785
QY 303 IENAKKCE--DFT--KKLEGEHAQLG--IENVTDENAKKAILITPAKDKGAAELEKL 355
DB 1786 DSKKKITELENDLITSKKKLEFTEKTKTSKFNLEER-----KDKELVKLNKE 1833
QY 356 FKAVENLAKAKEMLANVKEL 377
DB 1834 LELLKNDNSGAKKELSEKVSKL 1855
```

Search completed: March 18, 2002, 09:58:21  
Job time: 379 sec





Tue Mar 19 10:57:48 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds  
(without alignments)  
417.592 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850

Sequence: 1 MACNNNSGKDGTSANSADES.....AVENLAKAKEMLANSVKEL 377

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	51.1	210	2	G70218
2	938	50.7	210	2	S69927
3	876	47.4	193	2	S70287
4	863	46.6	191	2	I40153
5	853	46.1	177	2	I40129
6	826	44.6	178	2	I40125
7	823	44.5	178	2	I40104
8	774.5	41.9	211	2	I40145
9	773.5	41.8	193	2	S70279
10	739	39.9	178	2	I40124
11	709	38.3	210	2	I40144
12	695.5	37.6	211	2	S69918
13	692.5	37.4	211	2	I40277
14	690.5	37.3	193	2	S70280
15	686.5	37.1	211	2	I40268
16	686	37.1	190	2	S70273
17	685.5	37.1	209	2	I40273
18	683.5	36.9	191	2	S70278
19	675	36.5	194	2	S70277
20	674.5	36.5	209	2	S69917
21	670	36.2	212	2	S70254
22	666.5	36.0	193	2	S70276
23	664.5	35.9	193	2	S70274
24	664.5	35.9	211	2	I40278
25	664	35.9	212	2	S69922
26	662.5	35.8	191	2	S70284
27	662.5	35.8	193	2	S70265
28	662	35.8	194	2	S70268
29	661.5	35.8	211	2	S69932

30	657.5	35.5	209	2	I40142	outer surface prot
31	655	35.4	212	2	I40079	outer surface prot
32	650	35.1	212	2	I40143	outer surface prot
33	646	34.9	214	2	S69916	outer surface prot
34	641.5	34.7	193	2	S70286	outer surface prot
35	638	34.5	194	2	S70289	outer surface prot
36	635.5	34.4	191	2	S70288	outer surface prot
37	635.5	34.4	203	2	I40108	outer surface prot
38	631	34.1	212	2	S20543	outer surface prot
39	630.5	34.1	209	2	I40281	outer surface prot
40	630.5	34.1	209	2	I40285	outer surface prot
41	630	34.1	192	2	S70285	outer surface prot
42	628	33.9	210	2	S69925	outer surface prot
43	627	33.9	210	2	S69923	outer surface prot
44	627	33.9	212	2	S69921	outer surface prot
45	622.5	33.6	209	2	S69926	outer surface prot

## ALIGNMENTS

RESULT 1  
G70218  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence, revision 13-Feb-1998 #text\_change 21-Jul-2000  
C:Accession: G70218; I40269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; M01D:98065943  
A:Accession: G70218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <KLE>  
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu  
A:Reference number: I40269; M01D:96025162  
A:Accession: I40269  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:ID49497; NID:g707092; PIDN:BA08457.1; PID:g769684  
R:Jauris-Helpe, S.; Fuchs, R.; Moltz, W.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os  
A:Reference number: S37726; M01D:93268136  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:River, I.; Gibbs, C.P.; Schuster, R.; Dorrner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
A:Reference number: S70255; M01D:96296448  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIY>  
A:Cross-references: EMBL:I42887; NID:g858715; PIDN:AB36955.1; PID:g1695212  
A:Experimental source: strain Ip2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 51.1%; Score 945; DB 2; Length 210;  
 Best Local Similarity 97.9%; Pred. No. 5.5e-39;  
 Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNVLLAVKEVALLSSIDEIAAKA 60  
 DB 17 ISCNNSGKDGNTSANSADSVKGNLTETISKRTIDSNVLLAVKEVALLSSIDEIAAKA 76  
 QY 61 IGGKIHONNGLDTENNHNSILAGAVAI STLKOKLDGLNKGLEKIDAAKCSSEFTFN 120  
 DB 77 IGGKIHONNGLDTENNHNSILAGAVAI STLKOKLDGLNKGLEKIDAAKCSSEFTFN 136  
 QY 121 KIKKHTDGLKEGYTDADAKKAIKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVK 180  
 DB 137 KIKKHTDGLKEGYTDADAKKAIKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVK 196  
 QY 181 ELTSPVAESPCKP 194  
 DB 197 ELTSPVAESPCKP 210

RESULT 2  
 569927  
 outer surface protein C precursor - Lyme disease spirochete (strain PKa)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 A:Variety: strain PKa  
 C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
 C:Accession: S69927; S72659  
 R:Jauris-Heldke, S.; Ilegai, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.  
 J. Clin. Microbiol. 33, 1860-1866, 1995  
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia  
 A:Reference number: 140047; MUID:95395018  
 A:Accession: S69927  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-210 <JAU>  
 A:Cross-references: EMBL:X69589  
 R:Jauris, S.  
 A:Experimental source: Strain PKa  
 submitted to the EMBL Data Library, February 1994  
 A:Reference number: S72659  
 A:Accession: S72659  
 A:Molecule type: DNA  
 A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAM>  
 A:Cross-references: EMBL:X69589  
 C:Genetics:  
 A:Gene: ospC  
 Superfamily: Lyme disease spirochete surface protein C

Query Match 50.7%; Score 938; DB 2; Length 210;  
 Best Local Similarity 96.9%; Pred. No. 1.2e-38;  
 Matches 188; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNVLLAVKEVALLSSIDEIAAKA 60  
 DB 17 ISCNNSGKDGNTSANSADSVKGNLTETISKRTIDSNVLLAVKEVALLSSIDEIAAKA 76  
 QY 61 IGGKIHONNGLDTENNHNSILAGAVAI STLKOKLDGLNKGLEKIDAAKCSSEFTFN 120  
 DB 77 IGGKIHONNGLDTENNHNSILAGAVAI STLKOKLDGLNKGLEKIDAAKCSSEFTFN 136  
 QY 121 KIKKHTDGLKEGYTDADAKKAIKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVK 180  
 DB 137 KIKKHTDGLKEGYTDADAKKAIKANGTKTKGAEEIGKLFESVEVLSKAKEMLANSVK 196  
 QY 181 ELTSPVAESPCKP 194  
 DB 197 ELTSPVAESPCKP 210

RESULT 3  
 570287

outer surface protein C - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S70287  
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
 Mol. Microbiol. 18, 237-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70287  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-193 <LIY>  
 A:Cross-references: EMBL:442895; NID:9858723; PIDN:AAA37003.1; PID:91695220  
 A:Experimental source: strain 28354  
 C:Genetics:  
 A:Gene: ospC  
 Superfamily: Lyme disease spirochete surface protein C

Query Match 47.4%; Score 876; DB 2; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 1e-35;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 NSGKDGNTSANSADSVKGNLTETISKRTESNAVYLAVKEITLLASIDELATAIKG 257  
 DB 2 NSGKDGNTSANSADSVKGNLTETISKRTESNAVYLAVKEITLLASIDELATAIKG 61  
 QY 258 KIQONGGLAVAGHNGTLLAGAVTISKLITOKLDGLNKSKEKIDAAKCSSEFTTKL 317  
 DB 62 KIQONGGLAVAGHNGTLLAGAVTISKLITOKLDGLNKSKEKIDAAKCSSEFTTKL 121  
 QY 318 GEHAOLGIEENVTDENAKKAILITDAKDKGAEELEKLFRAVENLAKAKEMLANSVKEL 377  
 DB 122 GEHAOLGIEENVTDENAKKAILITDAKDKGAEELEKLFRAVENLAKAKEMLANSVKEL 181

RESULT 4  
 140153  
 outer surface protein C - Lyme disease spirochete (fragment)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 26-May-2000  
 C:Accession: I40153  
 R:Funng, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.  
 Infect. Immun. 62, 3213-3221, 1994  
 A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi  
 A:Reference number: I40153; MUID:94314437  
 A:Accession: I40153  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-191 <RES>  
 A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562  
 C:Genetics:  
 A:Gene: ospC  
 Superfamily: Lyme disease spirochete surface protein C

Query Match 46.6%; Score 863; DB 2; Length 191;  
 Best Local Similarity 99.4%; Pred. No. 4.2e-35;  
 Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 NSGKDGNTSANSADSVKGNLTETISKRTESNAVYLAVKEITLLASIDELATAIKG 258  
 DB 1 NSGKDGNTSANSADSVKGNLTETISKRTESNAVYLAVKEITLLASIDELATAIKG 60  
 QY 259 IQONGGLAVAGHNGTLLAGAVTISKLITOKLDGLNKSKEKIDAAKCSSEFTTKLE 318  
 DB 61 IQONGGLAVAGHNGTLLAGAVTISKLITOKLDGLNKSKEKIDAAKCSSEFTTKLE 120  
 QY 319 GEHAOLGIEENVTDENAKKAILITDAKDKGAEELEKLFRAVENLAKAKEMLANSVKEL 377  
 DB 121 GEHAOLGIEENVTDENAKKAILITDAKDKGAEELEKLFRAVENLAKAKEMLANSVKEL 179

RESULT 5  
140129  
outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40129; S54199  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40129  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-177 <RES>  
A:Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 46.1%; Score 853; DB 2; Length 177;  
Best Local Similarity 98.3%; Pred. No. 1.1e-34;  
Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 11 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIHONG 70  
|||||  
DB 1 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIHONG 60  
QY 71 LDFENHNGSLAGAYATISLITOKLDGLKNEGLKEKIDAKKSEFTFKLEKHTDLG 130  
|||||  
DB 61 LDFENHNGSLAGAYATISLITOKLDGLKNEGLKEKIDAKKSEFTFKLEKHTDLG 120  
QY 131 KEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSVKELTSPV 187  
|||||  
DB 121 KEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSVKELTSPV 177

RESULT 6  
140125  
outer surface protein C - Lyme disease spirochete (strain MDL) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40125; S54195  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40125  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-178 <RES>  
A:Cross-references: EMBL:X84779; NID:g793817; PIDN:CAA59250.1; PID:g793818  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 44.6%; Score 826; DB 2; Length 178;  
Best Local Similarity 98.8%; Pred. No. 2.3e-33;  
Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 205 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIQONG 264  
|||||  
DB 1 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIQONG 60  
QY 265 LAVEAGHNGTLGAYATISLITOKLDGLKNEGLKEKIDAKKSEDFPKLEGEHAOL 324  
|||||  
DB 61 LAVEAGHNGTLGAYATISLITOKLDGLKNEGLKEKIDAKKSEDFPKLEGEHAOL 120  
QY 325 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKRAKEMLANSVKEL 377  
|||||  
DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKRAKEMLANSVKEL 173

RESULT 7

I40104  
outer surface protein C - Lyme disease spirochete (strain 272) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40104; S54184  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40104  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-178 <RES>  
A:Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 44.5%; Score 823; DB 2; Length 178;  
Best Local Similarity 98.3%; Pred. No. 3.2e-33;  
Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 205 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIQONG 264  
|||||  
DB 1 NTSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAKAIGKKIQONG 60  
QY 265 LAVEAGHNGTLGAYATISLITOKLDGLKNEGLKEKIDAKKSEDFPKLEGEHAOL 324  
|||||  
DB 61 LAVEAGHNGTLGAYATISLITOKLDGLKNEGLKEKIDAKKSEDFPKLEGEHAOL 120  
QY 325 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKRAKEMLANSVKEL 377  
|||||  
DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKRAKEMLANSVKEL 173

RESULT 8  
I40145  
outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40145  
R:Stevenson, B.; Barthold, S.W.  
FEMS Microbiol. Lett. 124, 367-372, 1994  
A:Title: Expression and sequence of outer surface protein C among North American isol  
A:Reference number: I40143; MUID:95154673  
A:Accession: I40145  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-211 <RES>  
A:Cross-references: EMBL:U04282; NID:g2314881; PIDN:AA045540.1; PID:g434666  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.9%; Score 774.5; DB 2; Length 211;  
Best Local Similarity 81.0%; Pred. No. 8.1e-31;  
Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;  
QY 1 MACNNSGKDN-TSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAK 59  
:|||||  
DB 17 ISCNNSGKDN-TSANSADESVKGPNTLEIKTKITDSNAVLAVKEVEALLSIDETAK 76  
QY 60 AIGKKIHONGGLDENHNGSLAGAYATISLITOKLDGLKNEGLKEKIDAKKSEFT 119  
|||||  
DB 77 AIGKKIHONGGLDENHNGSLAGAYATISLITOKLDGLKNEGLKEKIDAKKSEFT 136  
QY 120 NKLEKHTDLGKEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSV 179  
|||||  
DB 137 NKLEKHTDLGKEGVTADAKKAILITDAKDKGAELKLFRAVENLAKRAKEMLANSV 196

QY 180 KELTSPVAESPKRP 194

Db 197 KELTSPPVAESPCKP 211

## RESULT 9

outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70279  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70279  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:U42898; NID:9858729; PIDN:AB37007.1; PID:91695223  
A:Experimental source: strain 25015  
C:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.8%; Score 773.5; DB 2; Length 193;  
Best Local Similarity 81.9%; Pred. No. 8.2e-31;  
Matches 158; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

QY 3 CNGSGKDGK-TSANSADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKAI 61  
Db 1 CNGSGKDGKMASTNPADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKAI 60  
QY 62 GKTHQNNGLDTENNHNSLLAGAYASTLIKOKLDGLKNEKKEIDAKKCEFTTK 121  
Db 61 GKTHQNNGLDTENNHNSLLAGAYASTLIKOKLDGLKNEKKEIDAKKCEFTTK 120  
QY 122 LKEKHTDLGKEGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAKEMLANSVKE 181  
Db 121 LKSHTELGKODDADAKKAILKTHNTKDKAEELDKLFEPVENSKAKEMLANSVKE 180  
QY 182 LTPSPVAESPCKP 194  
Db 181 LTPSPVAESPCKP 193

## ULT 10

outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40124; S54194  
R:Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40124  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-178 <RES>  
A:Cross-references: EMBL:X84782; NID:9793815; PIDN:CAA59253.1; PID:9793816  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 39.9%; Score 739; DB 2; Length 178;  
Best Local Similarity 89.0%; Pred. No. 3.3e-29;  
Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 205 NTSANSADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKAIKAGKTIQONGC 264  
Db 1 NTSANSADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKAIKAGKTIQONGC 60  
QY 265 LAVEAGHNGTLLAGAYITSLITQKLDGLKNSKLEKIEENAKKCSBDFTKLEGEHAOL 324

Db 61 LAVEAGHNGTLLAGAYITSLITQKIRWENSKIRKIEENAKKCSBDFTKLEGEHAOL 120  
QY 325 GTEWVDENNAKKAILLTDAAKKGAAELEKFAVENLAKAKEMLANSVKEL 377  
Db 121 GTEWVDENNAKKAILLTDAAKKGAAELEKFAVENLAKAKEMLANSVKEL 173

## RESULT 11

outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: S70282  
R:Stevenson, B.; Barthold, S.W.  
FEMS Microbiol. Lett. 124, 367-372, 1994  
A:Title: Expression and sequence of outer surface protein C among North American isol  
A:Reference number: I40143; MUID:95154673  
A:Accession: I40144  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: EMBL:U04281; NID:9434663; PIDN:AA03297.1; PID:9434664  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70282  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:U42893; NID:9858721; PIDN:AB37001.1; PID:91695218  
A:Experimental source: strain 297  
C:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 38.3%; Score 709; DB 2; Length 210;  
Best Local Similarity 75.4%; Pred. No. 1.1e-27;  
Matches 147; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNLITKINKITDSNAVLAKEVEALLSIDELAKA 75  
QY 61 IGKTHQNNGLDTENNHNSLLAGAYASTLIKOKLDGLKNEG-LKEKIDAKKCEFTT 119  
Db 76 IGKTHQNNGLDTENNHNSLLAGAYASTLIKOKLDGLKNEG-LKEKIDAKKCEFTT 118  
QY 120 NLEKHTDLGKEGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAKEMLANSV 179  
Db 136 AKLGEHTDLGKEGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAKEMLANSV 195  
QY 180 KELTSPPVAESPCKP 194  
Db 196 KELTSPPVAESPCKP 210

## RESULT 12

outer surface protein C precursor - Lyme disease spirochete (strain PBre)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
A:Variety: strain PBre  
C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69918; S72674; I40103  
R:Tauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69918  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA



```

Db      77  ||||| ||| | | :||||| ||| | | :| | | | :| | | | :| | | | :| |
      120  NKLKKEHTDLDGEGVTDADAKKAITLKANGTKKGAEELGKLFESYEVLSKAKEMLANSV 179
      137  KKLSDNOAELGIENXTDNDNAKAILKTHNAKDGAELVKLSESVAGLLKAAQAILANSV 196
Oy      180  KELTSPVVAESPCKP 194
      197  KELTSPVVAESPCKP 211
Db

```

Search completed: March 18, 2002, 09:56:47  
 Job time: 370 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:45 ; Search time 39.62 Seconds

(without alignments)  
348.880 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850  
Sequence: 1 MACNNSGKDGNTSANSADES.....AVENIAKAKEMIANSVKEL 377

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database : SwissProt\_39:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	51.1	210	1	OSCL_BORBU
2	631	34.1	212	1	OSCL_BORBU
3	397.5	21.5	214	1	VM24_BORHE
4	397	21.5	215	1	VM03_BORHE
5	174.5	9.4	1509	1	MYSN_ACACA
6	172	9.3	492	1	M5_STRPY
7	164.5	8.9	1164	1	BAG_STRPY
8	161.5	8.7	1790	1	USOL_YEAST
9	160	8.6	1433	1	REST_CHICK
10	159.5	8.6	1639	1	LMG1_DROME
11	159.5	8.6	1957	1	YD86_SCHPO
12	157	8.5	483	1	MG_STRPY
13	151.5	8.2	2017	1	MYSN_DROME
14	150	8.1	1938	1	MYSN_ACOIR
15	150	8.1	1966	1	MYSB_CAEEL
16	149	8.1	662	1	TLPA_BACSU
17	148	8.0	2116	1	MYG2_DICDI
18	148	8.0	3110	1	LMA2_HUMAN
19	147.5	8.0	1013	1	SCA4_RICRH
20	147.5	8.0	3210	1	CENF_HUMAN
21	147	7.9	564	1	M12_STRPY
22	147	7.9	1330	1	SMC3_YEAST
23	145	7.8	998	1	SCA4_RICAK
24	143	7.7	1093	1	TMF1_HUMAN
25	142.5	7.7	2869	1	RBP1_PLAVB
26	142	7.7	540	1	CH60_MYCPA
27	142	7.7	3672	1	LM2_CAEEL
28	141.5	7.6	875	1	TL21_YEAST
29	141.5	7.6	1786	1	LM1_MOUSE
30	141	7.6	1427	1	REST_HUMAN
31	140.5	7.6	1011	1	SCA4_RICAF
32	140.5	7.6	1786	1	LM1_HUMAN
33	140	7.6	1147	1	CGAL_HELPY

34	139.5	7.5	756	1	Y328_MYCSE
35	139.5	7.5	1935	1	MYSS_CYPCA
36	139	7.5	1391	1	MST2_DROHY
37	139	7.5	2349	1	TPR_HUMAN
38	138.5	7.5	944	1	NUF1_YEAST
39	138	7.5	724	1	HMWR_HUMAN
40	137.5	7.4	662	1	TLPB_BACSU
41	137	7.4	697	1	MFPI_LYCES
42	137	7.4	1500	1	SSP5_STRGN
43	137	7.4	137	1	MYHB_RABIT
44	136.5	7.4	473	1	YVCE_BACSU
45	136.5	7.4	1189	1	YJH6_YEAST

## ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD:	PRT:	210 AA.
AC	007337				
DT	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB19.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OG	Plasmid lp54.				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-9329332; PubMed-8098841;				
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,				
RA	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (Ospc) and the flagellin of Borrelia burgdorferi.;"				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-9329332; PubMed-8478108;				
RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant				
RT	major outer surface protein of Borrelia burgdorferi.;"				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-94041630; PubMed-8225587;				
RA	Padula S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.;"				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.;"				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-98065943; PubMed-9403685;				
RA	Frasar C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kariyave A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

EMBL	U01894	AA11208.1	-
DR	EMBL	AA16058.1	-
DR	EMBL	BA08457.1	-
DR	AE000792	AAC6329.1	-
DR	TIGR	BBB19	-
DR	InterPro	IPR001800	Lipoprotein_6
DR	Pfam	PF01441	Lipoprotein_6
DR	Prodom	PD001149	Lipoprotein_6
DR	PROSITE	PS00013	PROKARYOTIC LIPOPROTEIN
KW	Outer membrane	Lipoprotein	Signal; Plasmid; Antigen
KW	Complete proteome		
FT	SIGNAL	1	18
FT	CHAIN	19	210
FT	LIPID	19	19
SO	SEQUENCE	210 AA	22340 MW; 7A4FC97BF9177BF CRC64

QY	1	MACNNSGKRGKNTSANSADSEYVKGPMULTIINKITDSNAVLAVKEVEBALLSIDETIAANA	60
Db	17	ICONSNGKDGKNTSANSADSEYVKGPMULTEISKRTDSDNAVLAVKEVEBALLSIDETIAANA	76
QY	61	ISKRTIHQNNGLDPTENNHNHSGSLLAGAYASTLTKOKLGLKINEGKEKIDAAKCSSTFTN	120
Db	77	IGKRTIHQNNGLDPTENNHNHSGSLLAGAYASTLTKOKLGLKINEGKEKIDAAKCSSTFTN	136
Db	121	KIKKEKHTDLGKEGYTDADAKPAILKANGRTKGAEEIGKLFESVEYVLSRAAKEMLANSVK	180
Db	137	KIKKEKHTDLGKEGYTDADAKPAILKTNGTKTGKEEELGKLFESVEYVLSRAAKEMLANSVK	196
QY	181	ELTSPVPAESPKRP	194
Db	197	ELTSPVPAESPKRP	210

RESULT	2
ID	OSC2_BORBU
OSC2_BORBU	STANDARD;
OSC2_BORBU	PRT; 212 AA.
OSC2_BORBU	008137;
OSC2_BORBU	15-DEC-1998 (Rel. 37, Created)
OSC2_BORBU	15-DEC-1998 (Rel. 37, Last sequence update)
OSC2_BORBU	15-DEC-1998 (Rel. 37, Last annotation update)
OSC2_BORBU	OUTER SURFACE PROTEIN C PRECURSOR (PC).
OSC2_BORBU	OSPC.
OSC2_BORBU	Borrelia burgdorferi (Lyme disease spirochete),
OSC2_BORBU	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OSC2_BORBU	[1]_TaxId=139;
OSC2_BORBU	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
OSC2_BORBU	SRRIN-FRO;
OSC2_BORBU	MEDLINE=92219995; PubMed=1560779;
OSC2_BORBU	Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

DR	EMBL	Accession	Protein
DR	EMBL	A62162; CA444093.1; -	
DR	EMBL	X73624; CA452003.1; -	
DR	Interpro	IPR001800; Lipoprotein_6	
DR	Pfam	PF01441; Lipoprotein_6.1	
DR	ProDom	PD001149; Lipoprotein_6.1	
KW	Outer membrane; Lipoprotein; Signal; plasmid; Antigen.		
FT	SIGNAL	1	18
FT	CHAIN	19	212
FT	LIPID	19	19
SSQ	SEQUENCE	212 AA; 22499 MW; C206C231FEF2ETD4 CRC64;	

[illegible]

	RESULT	3
M24_BORHE		
D	VM24_BORHE	STANDARD;
C	P32778;	PRT;
T	01-OCT-1993 (Rel. 27, Created)	214 AA.
T	01-OCT-1993 (Rel. 27, Last sequence update)	
T	01-FEB-1994 (Rel. 28, Last annotation update)	
E	VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.	
N	VM24.	
S	Borrelia hermslii.	
G	Plasmid.	
C	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.	
X	NCBI_TaxID=140;	
N	[1]	
	SEQUENCE FROM N.A.	





CC LIGHT CHAIN SUBUNITS (MUC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MUC-2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS  
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE  
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.  
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING  
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Y00624; CA66863.1; -  
 CC PIR: A27224; A27224.  
 CC HSSP: P08799; 1MND.  
 CC InterPro: IPR000048; IQ.  
 CC InterPro: IPR001609; myosin\_head.  
 CC Pfam: PF00612; IQ; 2.  
 CC PRINTS: PR00063; myosin\_head; 1.  
 CC PRODOM: PD000355; MYOSINHEAVY.  
 CC SMART: SM00015; IQ; 1.  
 CC SMART: SM00242; MYSC; 1.  
 CC PROSITE: PS50096; IQ; 1.  
 CC MYOSIN: Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;  
 CC Methylation; Alkylation; Phosphorylation; Multigene family.  
 CC MYOSIN HEAD-LIKE.  
 CC  
 CC FT DOMAIN 1 789  
 CC FT DOMAIN 819  
 CC FT DOMAIN 790 819  
 CC FT DOMAIN 848 1509  
 CC FT DOMAIN 1226 1509  
 CC FT DOMAIN 1227 1509  
 CC FT DOMAIN 1253 1509  
 CC FT DOMAIN 1253 1509  
 CC FT DOMAIN 1482 1509  
 CC FT DOMAIN 1483 1509  
 CC FT NR\_BIND 182 189  
 CC FT DOMAIN 660 682  
 CC FT DOMAIN 766 780  
 CC FT MOD\_RES 133 133  
 CC FT MOD\_RES 700 700  
 CC FT MOD\_RES 1489 1489  
 CC FT MOD\_RES 1494 1494  
 CC FT MOD\_RES 1499 1499  
 CC FT MOD\_RES 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;  
 CC  
 CC SQ SEQUENCE

Query Match 9.4%; Score 174.5; DB 1; Length 1509;  
 Best Local Similarity 24.7%; Pred. No. 0.38;  
 Matches 98; Conservative 73; Mismatches 159; Indels 67; Gaps 19;

QY 28 EINKKITSNAVLAVKEVEA-----LSSIDEIAKAIGKIKHQ--ANGLDTENNHNSL 81  
 DB 934 ELQEEFSASNDLEQKRIKLEAKGELKASLEB--EENRRKALDEAKTVKSEENRELQDK 990  
 QY 82 LAGAVAVITLKOKLDKLGKNEGLKEKIDA---AKKSEFTFNKLEKKTDTGKGGVYDA- 137  
 DB 991 YEDEAAHDSLKKEEDLSRE-LRETKDALDAENISITLSKL--KNTGAGADVNNEL 1047  
 QY 138 -DAKFAIKANGTKKGAEELGKFESEYVSKAKELANSVKELTSPV----- 186  
 DB 1048 DDVATKTKQLKTKKSLKEELAQTRADLE- EKSGKEAASSKAKOLOQOEDARSEVDSL 1106  
 QY 187 -----VASEPKKPSWVNNSG-----KDGNTSANGADESVNG--PNLFEISKKIT-----ES 230  
 DB 1107 KSKLSAAEKSLKAKQDNRDLDQLEDERVYRANVDKOKKALEKLTLELDVYALDQK 1166

QY 231 NAVLAVKEIETLLASIDELATKAIGKIOONGIA--VEAGHNTLLAGATTISKLTQ 288  
 DB 1167 NA---AAQAOKTKTKTOYDE--TK---RRLDEAPASARLEKERNKALDEVAULTDLDLAE 1218  
 QY 289 KLGGLKXSEKLEKTI-----ENARK---CSDFPKLKEGEHAQIGINVTDENAKKAI 338  
 DB 1219 RDSGAQQRRLNIRISELSELENARKTKGTGASSEYVRLEGELERLNEBELLNOENRRAA- 1277  
 QY 339 LITDAKDKGAFLKLEFLFAVENLAKAKEMIANSVK 375  
 DB 1278 --AEKMLDKANLELELROEADDAARDNDKLVKDNKR 1312

RESULT 6  
 M5\_STRPY  
 ID M5\_STRPY STANDARD: PRT; 492 AA.  
 AC P02977;  
 DT 21-JUL-1986 (rel. 01, Created)  
 DT 01-FEB-1991 (rel. 17, Last sequence update)  
 DT 15-JUL-1999 (rel. 38, Last annotation update)  
 DE M PROTEIN, SEROTYPE 5 PRECURSOR.  
 GN EM5 OR SMP5.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 CC Streptococcus  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88186881; PubMed=3281944;  
 RA Miller L., Gray L., Beachey E., Kehoe M.;  
 RT "Antigenic variation among group A streptococcal M proteins.  
 RT Nucleotide sequence of the serotype 5 M protein gene and its  
 RT relationship with genes encoding types 6 and 24 M proteins.";  
 RL J. Biol. Chem. 263:5668-5673(1988).  
 RN [2]  
 RP SEQUENCE OF 43-212 AND 238-250.  
 RX MEDLINE=84162039; PubMed=6368549;  
 RA Manjula B.N., Acharya A.S., Mische S.M., Fairwell T., Fischetti V.A.;  
 RT "The complete amino acid sequence of a biologically active  
 RT 197-residue fragment of M protein isolated from type 5 group A  
 RT streptococci.";  
 RL J. Biol. Chem. 259:3686-3693(1984).  
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES  
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF  
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO  
 CC PHAGOCYTOSIS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.  
 CC -1- SIMILARITY: TO OTHER M PROTEINS.  
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS  
 CC IN THE REGION OF THE MEMBRANE ANCHOR.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M20374; AAA6976.1; -  
 CC PIR: A03501; M5OMP.  
 CC PIR: A28616; A28616.  
 CC InterPro: IPR001899; Gram\_pos\_anchor.  
 CC InterPro: IPR003345; M\_repeat.  
 CC Pfam: PF00746; Gram\_pos\_anchor; 1.  
 CC Pfam: PF02370; M\_9  
 CC PRINTS: PR00015; GP0SANCHOR.  
 CC PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
 CC Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
 CC Transmembrane; Coiled coil; Signal.  
 CC SIGNAL 1 42  
 CC FT CHAIN 43 492 M PROTEIN, SEROTYPE 5.  
 CC FT DOMAIN 43 466 EXTRACELLULAR (POTENTIAL).



QY 345 KDKGAELERKFAVENL 362  
Db 429 NENNOKIE-LTVSPENI 445

RESULT 8  
USOL\_YEAST  
ID USOL\_YEAST STANDARD: PRT: 1790 AA.  
AC P25386;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.  
GN USOL OR INT1 OR YDL058W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces;  
NCBI\_TaxID=4932;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=X2180-1A;  
RX MEDLINE=91185402; PubMed=2010462;  
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
Yamasaki M.;  
RT "A cytoskeleton-related gene, usol, is required for intracellular  
protein transport in Saccharomyces cerevisiae.";  
RT J. Cell Biol. 113:245-260(1991).  
RN [2]  
RP SEQUENCE OF 782-1790 FROM N.A.  
RA Hostetter M.K., Herman D.J., Bendel C.M., McCellan M., Tao N.,  
Kendrick K.E.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-8 FROM N.A.  
RA Bai Y., Symington L.S.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
COMPLEX.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
ER AND THE GOLGI COMPLEX.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
CC -1- SIMILARITY: BELONGS TO THE VDE/USOL/YBL047C FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: X54378; CAA38253.1;  
DR EMBL: L03188; AAB00143.1;  
DR EMBL: 053668; AAB66559.1;  
DR PIR: A38455; A38455.  
DR HSSP: P80220; IDIP.  
DR SGD: S0002216; USOL.  
DR InterPro: IPR002017; Spectrin.  
KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
FT DOMAIN 1 724 GLOBULAR HEAD.  
FT DOMAIN 725 1790 COILED COIL (POTENTIAL).  
FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).  
FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.  
FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).  
FT CONFLICT 847 847 G -> E (IN REF. 2).  
FT CONFLICT 924 924 E -> K (IN REF. 2).  
FT CONFLICT 1253 1253 V -> I (IN REF. 2).  
FT CONFLICT 1319 1319 I -> V (IN REF. 2).  
FT CONFLICT 1461 1461 N -> S (IN REF. 2).

FT CONFLICT 1581 1581 G -> S (IN REF. 2).  
FT CONFLICT 1600 1600 I -> V (IN REF. 2).  
FT CONFLICT 1661 1661 R -> S (IN REF. 2).  
FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).  
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 8.7%; Score 161.5; DB 1; Length 1790;  
Best Local Similarity 21.7%; Pred. No. 1.8;  
Matches 99; Conservative 72; Mismatches 178; Indels 107; Gaps 16;

QY 3 CNGSGKDG-----NTSANSADSEYKGNLLEIKKITTDSNAYLAVKEVALSS 52  
Db 929 CNNLSKEKEHISKELVEKRSFOSHDNLV--AKTEKIKLSANNYKMOQA--ENESLKA 984  
QY 53 IDEIAKKAIGKRIHONNGID--TENNNGSLAGAY-----AISTL-----TKOR 95  
Db 985 VEESKNESSTIQLSNLONKIDSMGKFNFOIERGSIENKIDQLKTTISDEQTKREILSK 1044  
QY 96 LDGLKNEG-----LKEKIDAKKCSFTFKLKE-----KHTDLGKGGVD 136  
Db 1045 SSSSDEYESQISILKEKLEATTATANDENVAKISELTRTRELEALAVKKNLKELETK 1104  
QY 137 ADAPKALIK-----ANGTKTK-----GAEELGKLESEYEVLSKAA 171  
Db 1105 LETSEKALKEVENEHLEKEKIQLEKREATEETQOQNSLANSLESEKEHEDLAOLKRY 1164  
QY 172 KEMLANSVKELTSPV--VAESPCKPSMVNNSGDKDQNTSANSADSEYKGNLLEISKITE 229  
Db 1165 EEOJANNEQVNEITSOINDEITSTOENESIKRKDELEGEVKAKRSTSEOSNLKSE 1224  
QY 230 SNAVYLAVKETE-----TLASIDELATKAIGKIIQONGLAIVEAHNNTLLAGAT 281  
Db 1225 IDALNLIQIKELKKNETNEASLESISVSESYV--KINE--LDQCNFK----- 1270  
QY 282 ISKLTITKLGKLNSEKLEKIKINAKKCEDPTFKKIEGSHAOLGIENVTDENAKKAILIT 341  
Db 1271 -EKEVSELEKLAISEDKSKSTIETLOKESEKIEKIEDATATYETLKIQLKRTWLSK----- 1324  
QY 342 DAAKDKGAELERKFAVENLAKAAKEMLANSVKEL 377  
Db 1325 --AKESSESLSLKTKTSSEERKNAEDELKLNKEI 1358  
RESULT 9  
REST CHICK  
ID REST CHICK STANDARD: PRT: 1433 AA.  
AC 042184; 042228; 057563; 057564;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).  
GN RSN.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98137792; PubMed=9469933;  
RA Griparic L., Volosky J.M., Keller T.C. III;  
RT "Cloning and expression of chicken CLIP-170 and restin isoforms.";  
RT Gene 206:195-208(1998).  
RN [2]  
RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).  
RC TISSUE=Pectoralis muscle;  
RA Griparic L., Keller T.C. III;  
RT "Identification and expression of two novel CLIP-170/Restin isoforms  
expressed predominantly in muscle.";  
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RL -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN  
THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE  
 CC CYTOSKELETON (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF014012; AAC60344.1; -;  
 DR EMBL: AF020764; AAC60345.1; -;  
 DR EMBL: AF045650; AAC03547.1; -;  
 DR EMBL: AF045651; AAC03548.1; -;  
 DR InterPro: IPR000938; CAP-GLY;  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF01302; CAP-GLY; 2.  
 DR SMART: SM00343; Znf\_C2HC; 1.  
 DR PROSITE: PS00845; CAP-GLY\_1; 2.  
 KM Cytokeleton; Microtubules; Colled coil; Alternative splicing.  
 FT DOMAIN 79 121  
 FT DOMAIN 144 207  
 FT DOMAIN 207 277  
 FT DOMAIN 235 332  
 FT DOMAIN 351 1353  
 FT DOMAIN 1414 1427  
 FT VARSPLIC 458 492  
 FT VARSPLIC 458 492  
 FT VARSPLIC 803 803  
 FT VARSPLIC 458 458  
 FT CONFLICT 309 309  
 FT CONFLICT 440 440  
 SQ SEQUENCE 1433 AA; 161026 MW; 5631CB8683458E23 CRC64;  
 Query Match 8.6%; Score 160; DB 1; Length 1433;  
 Best Local Similarity 25.9%; Pred. No. 1.6; Indels 58; Gaps 17;  
 Matches 96; Conservative 58; Mismatches 158; Indels 58; Gaps 17;  
 QY 28 EINKKTTDSNAVLAVKEVEALLSSIDEIAKAIIGKRIHONNGIDTENNHGSLLAGAYA 87  
 DB 524 ESSKHIDVDVTSILLOEI-----SILQEKMA-AAGKE-HQREMSLKEFESSEALRKE 577  
 QY 88 ISLT-IKQKIDGLKNEIKKIDAKR-----CSEFTFNKIKE---KITDLGKE----- 132  
 DB 578 IKLSASNEKMGKENEKLRKRLHANKENSVDTELMKSKLESALASHQAAMEELKVSFNK 637  
 QY 133 GYTDADAKKAILKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVKEITSPVAESPK 192  
 DB 638 GVGAGTAEFELTOMKVKALDE--NEMSNIKIKOENESOHKLELAKKLEETEE 695  
 QY 193 KPSMVNNSGKGDNSTANSADSVKGNLPTEISKRTESNAVLAKEIEFLASIDELAT 252  
 DB 696 KEOTLEF-----LKAKLESVEDQHLVEMEDTLNQLQAEIKVKELDLVLAQACNE-QT 746  
 QY 253 KAIGKKIQONGGLAVENGHGTL-----LAGAYTSKILITQKL-DGLKNSKRIKIKINAK 307  
 DB 747 KILGSLTQO-----IRSEKSLDLALOKANSEKLEIQSLQLOAAEQIONLETER 801  
 QY 308 KCEDEFTKKEGHAQIGIENVTDENAKKAILITDAKDGCAEL-----EKLFAV---E 360  
 DB 802 --VSNTLKEIQGKEQKL-----LDLEKMLSANOVKDSLEKELDLKKEFTSAVVGAE 852  
 QY 361 NLAKAKEML 370  
 DB 853 NAQRAQOETI 862

RESULT 10  
 ID LMGI\_PROME STANDARD; PRT; 1639 AA.  
 AC P15215; 024373; Q9VT18;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 20-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).  
 GN LANB2 OR LAMC1 OR LAMG1 OR CG3322.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S, AND OREGON-R;  
 RX MEDLINE=91299161; PubMed=1840513;  
 RA Chl H.-C., Jumiaga D., Wang S.Y., Hui C.-F.;  
 RT "Structure of the Drosophila gene for the laminin B2 chain.";  
 RL DNA Cell Biol. 10:451-466(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R;  
 RX MEDLINE=89109164; PubMed=2912972;  
 RA Chl H.-C., Hui C.-F.;  
 RT "Primary structure of the Drosophila laminin B2 chain and comparison  
 RT with human, mouse, and Drosophila laminin B1 and B2 chains.";  
 RL J. Biol. Chem. 264:1543-1550(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90037237; PubMed=2808533;  
 RA Montell D.J., Goodman C.S.;  
 RT "Drosophila laminin: sequence of B2 subunit and expression of all  
 RT three subunits during embryogenesis.";  
 RL J. Cell Biol. 109:2441-2453(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintsov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Buritis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferrara C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RP SEQUENCE OF 344-1639 FROM N.A.  
 RC STRAIN-OREGON-R;  
 RX MEDLINE=8830364; PubMed=3405777;  
 RA Chl H.-C., Hui C.-P.;  
 RT "cDNA and amino acid sequences of *Drosophila* laminin B2 chain.";  
 RL Nucleic Acids Res. 16:7205-7205(1988).  
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC -1- SUBCELLULAR LOCATION: BASEMENT MEMBRANES (MAJOR COMPONENT).  
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- SIMILARITY: CONTAINS VI AND IV ARE GLOBULAR.  
 CC -1- SIMILARITY: CONTAINS I LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -1- SIMILARITY: CONTAINS I LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS I LAMININ DOMAIN IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M58417; AAA28665.1; -;  
 DR EMBL; M25063; AAA28664.1; -;  
 DR EMBL; AE003551; AAF50238.1; -;  
 DR EMBL; X07806; CAA30665.1; -;  
 DR PIR; A31483; MMEFB2.  
 DR HSP; P02468; 1TLE.  
 DR FLYBASE; FBgn0002528; Lamb2.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001886; LamNT.  
 DR InterPro; IPR000034; Laminin\_B.  
 DR Pfam; PF00052; Laminin\_B\_1.  
 DR Pfam; PF00053; Laminin\_EGF\_10.  
 DR ProDom; PD002082; LamNT; 1.  
 DR ProDom; PD003031; Laminin\_B\_1.  
 DR SMART; SM00180; EGF\_Lam; 10.  
 DR SMART; SM00281; EGF\_Like; 1.  
 DR SMART; SM00136; LamNT; 1.  
 DR SMART; PS00022; EGF\_1; 8.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 11.  
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 DR Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 KW LAMININ EGF-LIKE DOMAIN; Cell adhesion; Repeat; Signal.  
 FT SIGNAL 1 33  
 FT CHAIN 34 1639 LAMININ GAMMA-1 CHAIN.  
 FT DOMAIN 34 298 LAMININ N-TERMINAL (DOMAIN VI).  
 FT DOMAIN 299 523 4.5 x LAMININ EGF-LIKE REPEATS (DOMAIN  
 FT V).  
 FT DOMAIN 299 358 LAMININ EGF-LIKE 1.  
 FT DOMAIN 359 413 LAMININ EGF-LIKE 2.  
 FT DOMAIN 414 460 LAMININ EGF-LIKE 3.  
 FT DOMAIN 461 513 LAMININ EGF-LIKE 4.  
 FT DOMAIN 514 523 LAMININ EGF-LIKE 5 (N-TERMINAL).  
 FT DOMAIN 524 709 LAMININ DOMAIN IV.

FT	DOMAIN	710	1049	6.5 x LAMININ EGF-LIKE REPEATS (DOMAIN III).
FT	DOMAIN	710	743	LAMININ EGF-LIKE 5 (C-TERMINAL).
FT	DOMAIN	744	792	LAMININ EGF-LIKE 6.
FT	DOMAIN	793	846	LAMININ EGF-LIKE 7.
FT	DOMAIN	847	901	LAMININ EGF-LIKE 8.
FT	DOMAIN	902	955	LAMININ EGF-LIKE 9.
FT	DOMAIN	956	1003	LAMININ EGF-LIKE 10.
FT	DOMAIN	1004	1049	LAMININ EGF-LIKE 11.
FT	DOMAIN	1050	1609	LAMININ II AND I.
FT	DOMAIN	1087	1109	COILED COIL (POTENTIAL).
FT	DOMAIN	1144	1247	COILED COIL (POTENTIAL).
FT	DOMAIN	1306	1627	COILED COIL (POTENTIAL).
FT	DISULFID	299	308	BY SIMILARITY.
FT	DISULFID	301	322	BY SIMILARITY.
FT	DISULFID	324	333	BY SIMILARITY.
FT	DISULFID	336	356	BY SIMILARITY.
FT	DISULFID	359	384	BY SIMILARITY.
FT	DISULFID	361	384	BY SIMILARITY.
FT	DISULFID	387	396	BY SIMILARITY.
FT	DISULFID	399	411	BY SIMILARITY.
FT	DISULFID	414	426	BY SIMILARITY.
FT	DISULFID	416	432	BY SIMILARITY.
FT	DISULFID	434	443	BY SIMILARITY.
FT	DISULFID	446	458	BY SIMILARITY.
FT	DISULFID	461	475	BY SIMILARITY.
FT	DISULFID	463	482	BY SIMILARITY.
FT	DISULFID	484	493	BY SIMILARITY.
FT	DISULFID	496	511	BY SIMILARITY.
FT	DISULFID	744	753	BY SIMILARITY.
FT	DISULFID	746	760	BY SIMILARITY.
FT	DISULFID	762	771	BY SIMILARITY.
FT	DISULFID	774	790	BY SIMILARITY.
FT	DISULFID	793	801	BY SIMILARITY.
FT	DISULFID	795	811	BY SIMILARITY.
FT	DISULFID	814	823	BY SIMILARITY.
FT	DISULFID	826	844	BY SIMILARITY.
FT	DISULFID	847	861	BY SIMILARITY.
FT	DISULFID	849	868	BY SIMILARITY.
FT	DISULFID	871	880	BY SIMILARITY.
FT	DISULFID	883	899	BY SIMILARITY.
FT	DISULFID	902	919	BY SIMILARITY.
FT	DISULFID	904	926	BY SIMILARITY.
FT	DISULFID	928	937	BY SIMILARITY.
FT	DISULFID	940	953	BY SIMILARITY.
FT	DISULFID	956	968	BY SIMILARITY.
FT	DISULFID	958	975	BY SIMILARITY.
FT	DISULFID	977	986	BY SIMILARITY.
FT	DISULFID	989	1001	BY SIMILARITY.
FT	DISULFID	1004	1016	BY SIMILARITY.
FT	DISULFID	1006	1022	BY SIMILARITY.
FT	DISULFID	1024	1033	BY SIMILARITY.
FT	DISULFID	1036	1047	BY SIMILARITY.
FT	DISULFID	1050	1050	INTERCHAIN (PROBABLE).
FT	DISULFID	1053	1053	INTERCHAIN (PROBABLE).
FT	DISULFID	1631	1631	INTERCHAIN (PROBABLE).
FT	CARBOHYD	147	147	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 8.6%; Score 159.5; DB 1; Length 1639;  
 Best Local Similarity 23.9%; Pred. No. 2;  
 Matches 105; Conservative 56; Mismatches 162; Indels 115; Gaps 19;

QY 2 ACNNGKGTGNTSANSAD-----ESVGPNTLEIKRTDSNVVLAVKEVALLSSIDE 55  
 DB 1107 ADNNGGDSGGTYAEVITDLHKHLDVSR-EHLVSADKQQAANEIDRAKQNTYLLQITE 1165  
 QY 56 IAAKAIKKIKHQNNGDLPENNNHNSLLAGAVYSTLKQALDKNKGKTKIDARKCS 115  
 DB 1166 NAKKEL-----QALDLNDGGAQALAKRAKESVBERGQ-----SEQISDISREA 1210  
 QY 116 EFTYNNKIK-EKHIDLGKGVYTDADAKAALIKAN-----GTRTKG--AEELGK 159  
 DB 1211 RALADKLESAQFDLKNA-----KDARKDAVEKAHQKLAKSAIDQLKIGTELRSVGLLEISH 1266

```

OY 160 LESEVLSKAKE-----MLANSVELTSPVVAESPKPSMVNNSGKDCNT 206
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1267 VKOSLGTVVOTSKRALKANVEYDTALTLLNDVNRQTOPIDISOLK-----KDAVA 1318
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 207 SANSADSVGPNLITEISKITESNAVLAVKEIETLLASIDELATKAIGKKIOQ-----261
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1319 ANRADBLK--QITELS-----NSNGELFADFEPEQEL-----TALAKRAEQOOL 1365
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 262 -----NGLAVERGHNGTLLGAVYTSKITOKIDGKINSE-----KL 299
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1366 IELLERAKAHDRAKATKAVEGODN-TLEKANNTEYKLAGFQSDYQSSSESAEKALQTVPMI 1424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 300 KEKIEVAKKSEDPFTKLEBEHNOIGENTYDENAKAIL--ITDAKDKGAELEKLF 357
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1425 EKEIÖNNEISIAOEALDANKN--ANPAKNAOEALKVAEQASKD-----AELIR 1475
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 358 AVENLAKAAKEMLANSVKEL 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1476 RKANETKVAARNLREADOL 1495
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      RESULT 11
      YD86_SCHPO STANDARD: PRT: 1957 AA.
      AC 010411:
      DT 01-OCT-1996 (Rel. 34, Created)
      DT 01-OCT-1996 (Rel. 34, Last sequence update)
      DT 01-OCT-1996 (Rel. 34, Last annotation update)
      DE HYPOHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
      OS SPAC1F3.06C.
      OC Schizosaccharomyces pombe (Fission yeast).
      OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
      OC Schizosaccharomycetales; Schizosaccharomycetaceae;
      OC Schizosaccharomycetes.
      OX NCBI_TaxID=4896;
      [1]
      RN SEQUENCE FROM N.A.
      RC STRAIN=972;
      RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
      RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
      CC -----
      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
      CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
      CC the European Bioinformatics Institute. There are no restrictions on its
      CC use by non-profit institutions as long as its content is in no way
      CC modified and this statement is not removed. Usage by and for commercial
      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
      CC or send an email to license@isb-sib.ch).
      -----
      EMBL: 270690; CAA94624.1; -.
      NM Hypothetical protein.
      SO SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

```

```

DB 534 INELKGLQTEINSEHLSLSLSTLAKEAAVATN-----NELSEKNSLQTLGNF 586
OY 172 KEMIANVKEITSPVVAESPKPSMVNNSGKDCNTSANSADSVKGNLITEISKITESN 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 587 OEKLAKSVOL-----KENQONSSLDTSFK--KLNSHOELENNH 625
OY 232 AVYLAVEIEFILLASIDELATKAIGKKIOONGGLAVAGHNGTLLAGAVYTSKITOK--289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 626 QTI--TQOLDKTSKLOQLERANFE-QKESTLSDENDRLRTLKLEESNKSILINKOE 682
OY 290 -LDGL-KNSEKLK--KIENAKKSE-----DFTKLEGEHAOLGIE-----NVTD 331
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 683 DVOSLEKNIQTLKEDLKSEALRFSLEAKNLEVIDNLKGRKETELEAQRNDHSLSD 742
OY 332 ENAKKAILITDAADKGALEKLFKAVENLAKAKEN-----LANSVEL 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 743 AKNTNAILSELT--KSEEDVKRLTAVETLTODSKAMKOSFTSLVNSYOSI 792
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      RESULT 12
      M6_STRPY STANDARD: PRT: 483 AA.
      ID M6_STRPY
      AC P08089;
      DT 01-AUG-1988 (Rel. 08, Created)
      DT 01-AUG-1988 (Rel. 08, Last sequence update)
      DT 20-AUG-2001 (Rel. 40, Last annotation update)
      DE M PROTEIN, SEROTYPE 6 PRECURSOR.
      GN EMM6.
      OS Streptococcus pyogenes.
      OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
      OC Streptococcus.
      OX NCBI_TaxID=1314;
      [1]
      RN SEQUENCE FROM N.A.
      RX MEDLINE=86111835; PubMed=3511046;
      RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
      RL "Relationship of M protein genes in group A streptococci."
      CC Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
      CC - FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
      CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
      CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
      CC PHAGOCYTOSIS.
      CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
      CC - SIMILARITY: TO OTHER M PROTEINS.
      CC - SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
      CC IN THE REGION OF THE MEMBRANE ANCHOR.
      CC -----
      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
      CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
      CC the European Bioinformatics Institute. There are no restrictions on its
      CC use by non-profit institutions as long as its content is in no way
      CC modified and this statement is not removed. Usage by and for commercial
      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
      CC or send an email to license@isb-sib.ch).
      -----
      EMBL: M1138; AAA26920.1; -.
      DR PIR: A26297; A26297.
      DR InterPro: IPR001899; Gram_pos_anchor.
      DR InterPro: IPR003345; M_repeat.
      DR Pfam: PF02370; M_9.
      DR PRINTS: PR00015; GPOSANCHOR.
      DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
      DR Virulence: Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
      KW Transmembrane; Coiled coil; Signal.
      FT SIGNAL 1 42

```







OY 316 KLEGEHAOLGIENVTDENAKKAILITD-----AAKDKGALEK-LFRAVENL- 362  
 DB 1587 ELKAOHEELEDQLEEDAKLELEVNMKQALRSOPFERDLAKREGAEKRGVLKQKDLDE 1646  
 OY 363 -----AKAKEMLANSVKEL 377  
 DB 1647 TELDEERKORTAFAVASKKKEGLKEI 1673  
 RESULT 14  
 MYS\_AEOIR STANDARD; PRT; 1938 AA.  
 AC P24733;  
 DT 01-MAR-1992 (rel. 21, last sequence update)  
 DT 20-AUG-2001 (rel. 40, last annotation update)  
 DE MYOSIN HEAVY CHAIN, STRIATED MUSCLE.  
 OS Aequipecten irradians (Bay scallop).  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea; Pectinoidae; Pectinidae; Argopecten.  
 NC NCBI\_TaxID=31199;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adductor muscle;  
 RX MEDLINE=92011595; PubMed=1917970;  
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;  
 RT Complete primary structure of a scallop striated muscle myosin heavy chain. Sequence comparison with other heavy chains reveals regions that might be critical for regulation."  
 RL J. Biol. Chem. 266:18469-18476(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adductor muscle;  
 RX MEDLINE=94119133; PubMed=8127365;  
 RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalborkis V.N., Szent-Gyorgyi A.G., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2.8-A resolution."  
 RL Nature 368:306-312(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.  
 RX MEDLINE=96419133; PubMed=8805510;  
 RA Houdusse A., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2-A resolution: implications for regulation."  
 RL Structure 4:21-32(1996).  
 RN [1]  
 RP FUNCTION: MUSCLE CONTRACTION.  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL, X55714; CAA39247.1; -

DR PIR; S13557; S13557.  
 DR PIR; A40997; A40997.  
 DR PDB; 1SCM; 30-APR-94.  
 DR PDB; 1WDC; 11-JUL-96.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00663; myosin\_head; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR ProDom; PD000355; myosin\_head; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR SMART; PS50096; IQ; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.  
 FT DOMAIN 1 777  
 FT DOMAIN 778 805  
 FT DOMAIN 836 1938  
 FT DOMAIN 836 1938  
 FT NE\_BIND 176 183  
 FT MOD\_RES 693 693  
 FT MOD\_RES 703 703  
 FT HELIX 778 821  
 FT TURN 822 823  
 FT TURN 825 833  
 FT TURN 834 835  
 SO SEQUENCE 1938 AA; 22821 MW; ASCCEA127D1A4896 CRC64;

Query Match 8.1%; Score 150; DB 1; Length 1938;

Best Local Similarity 20.4%; Pred. No. 6.4; Matches 81; Conservative 65; Mismatches 166; Indels 86; Gaps 12;

OY 18 DESKQGNLEIRNKINDSNAVLAVKEVEALLSSIDEIAKATGKTHONGLDT---- 73  
 DB 925 DEEDDAADLEGIRKKMKADNA-----NLKDDIDLENTLOKAODRAHKNOISTLOGE 978  
 OY 74 ---ENNHSILAGAVAISTLIKOKLDLKNGLKRIKIDAKKSETFTNKKKHTDLG 130  
 DB 979 ISQDDEHIGLNKKRKALKEANKKTSLSLOAE--EDCNHNLKIKALKEQADLEENLLE 1036  
 OY 131 KEGYTDAAKEAIIKANGTYTKGAEEIGKLFESYEVLSKAKEMLANSVKELTSPVAES 190  
 DB 1037 REKKVRGDVERA-----RKRVEDDLKSTOENVEDLERVARELEN-----VRK 1080  
 OY 191 PKRPSMNNNGSKGNTSANSADSEVK--GRNLTETISKRI--TESNAVYLVAKETITLASI 247  
 DB 1081 EAEISSLSNLSKLEEDQNLVSOLOKRIKELQARIEELELEERNAKRAKVEQRAELREL 1140  
 OY 248 DELATKAIIGKKIOONG-----LAVENGHTLLAGAYTISKL----- 285  
 DB 1141 EEL-----GERLDEAGATSTAOELNKKREAEELIKIRDLDEASLQHEAQSALRKKHQD 1195  
 OY 286 -ITQKLDGLNKKSEKLEKRIENAKK-----CSBDFPKLGEHAQ 323  
 DB 1196 AANEMADQVQLOQVKSKELEKDKDLKREMDLESQTHNNKKNKGCSEKVKQFESQMSD 1255  
 OY 324 LGIENVTDENAKKAILITDAKDKGALEKLFRAVEN 361  
 DB 1256 L--NARLEDSORSINELQSOKSRLQAEENSIDLTRQLED 1290  
 RESULT 15  
 MYS\_CAEEL STANDARD; PRT; 1966 AA.  
 AC P02566;  
 DT 21-JUL-1986 (rel. 01, last sequence update)  
 DT 21-JUL-1986 (rel. 01, last sequence update)  
 DT 20-AUG-2001 (rel. 40, last annotation update)  
 DE MYOSIN HEAVY CHAIN B (MHC B).

GN UNC-54 OR MYO-4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 ON NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83273600; PubMed=6576334;  
 RA Karn J., Brenner S., Barnett L.;  
 RT "Protein structural domains in the Caenorhabditis elegans unc-54  
 myosin heavy chain gene are not separated by introns.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).  
 RP SEQUENCE OF 850-1966 FROM N.A.  
 RX MEDLINE=82272395; PubMed=7202124;  
 RA McLachlan A.D., Karn J.;  
 RT "Periodic charge distributions in the myosin rod amino acid sequence  
 match cross-bridge specings in muscle.";  
 RL Nature 299:226-231(1982).  
 RN (3)  
 RP SEQUENCE OF 1876-1966 FROM N.A.  
 RX MEDLINE=83232892; PubMed=6571695;  
 RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,  
 RA Waterston R.H.;  
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber  
 nonsense mutations via altered transfer RNA.";  
 RL Cell 33:575-583(1983).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS  
 PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE  
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.  
 ELEGANS.  
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY  
 WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: J01050; AAA28124.1; -  
 DR EMBL: V01494; CAA24738.1; -  
 DR PIR: A02992; MWKW.  
 DR HSSP: P08799; IMND.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR Pfam: PF01576; myosin\_tail.1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head.1.  
 DR SMART: SM00242; MYSC.1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Multigene family.  
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.  
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).  
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).  
 FT DOMAIN 1165 1176 HINGE.  
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMW).

FT	NP_BIND	177	184	ATP (BY SIMILARITY).
FT	DOMAIN	665	687	ACTIN-BINDING.
FT	DOMAIN	769	783	ACTIN-BINDING.
FT	MOD_RES	128	128	METHYLATION (TRI-) (POTENTIAL).
FT	MOD_RES	705	705	ALKYLATION (SH-1).
FT	MOD_RES	715	715	ALKYLATION (SH-2).
FT	CONFLICT	1337	1337	E -> R (IN REF. 2).
FT	CONFLICT	1880	1880	I -> L (IN REF. 2).
SO	SEQUENCE	1966 AA;	225125 MW;	B6F0BB2FE27B67F CRC64;

Query Match 8.1%; Score 150; DB 1; Length 1966;  
 Best Local Similarity 22.7%; Pred. No. 6.5;  
 Matches 90; Conservative 76; Mismatches 149; Indels 82; Gaps 20;

QY	16	SADSVKGPNTLEIKKTIKIDSNVAVLAVK---EVALLSIDET--AKAIGKTIHQ--NN 69
DB	1358	STEEIEGKN--ELIRQSKANADIQMKARPEGGLKA-DELDARROAKINELQ 1414
QY	70	GLDTENNNGSL-LAGAYAIITLIKIDGLKNEGLKEKIDAKKCEFTNKLKEKHTD 128
DB	1415	ALDANKSKNASLEKTKSRVLGDDAQYDVERANGVASALEKKKGPKIIDEKAKTDD 1474
QY	129	LKEGVTDADAKAELTKANGT---KTGADELKLFESYEVLSKAKEMLANSVKELTSP 185
DB	1475	LAAE---LDGAQRDLKNTSTDLFAKNAQE--ELAEVVELIRREKNS-LSQETIKDLTDQ 1527
QY	186	VVAESPKRPSVNVNNGSKDNTSANSADSVKGPNL--TEISKKTIESNAVLAVKEIETL 243
DB	1528	L-----GEGRSVHEMOKIIRLEIEKEKELQHALDENVA-ALAESESKVL 1571
QY	244	LASIDELATKA-IGKTIQ-----NGLAVDAGHGTLLAGAVTIS 283
DB	1572	RAQVEVSQIRSEIEKRIQKEEKEEEMTKRNHARALESQSLFTEAKGAKELLR----- 1625
QY	284	KLIQKIDGLKNEKTKKIEKNKAKSCSDPTKKEGHAQGINVDENAKKAILITDA 343
DB	1626	--IKKLEGG--DINLELALDHAKKANADAKKNLR-----YQOVRELOLQVEE 1671
QY	344	AKDKGALEKLFKAVE--NLAKAKE--MLANSVKE 376
DB	1672	EORNGADTRQOFNAEKRRATLQSEKEBELLVANPAE 1708

Search completed: March 18, 2002, 10:11:49  
 Job time: 967 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:53 ; Search time 124.19 Seconds  
(without alignments)  
444.035 Million cell updates/sec

Title: US-09-596-746a-30

Sequence: 1850  
1 MACNSGKDGNTSANSADSE.....AVENLAKAKEMLANSVKEL 377

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: SP\_Archea:\*  
2: SP\_Bacteria:\*  
3: SP\_Fungi:\*  
4: SP\_Human:\*  
5: SP\_Invertebrate:\*  
6: SP\_Mammal:\*  
7: SP\_Mhc:\*  
8: SP\_Organelle:\*  
9: SP\_Phage:\*  
10: SP\_Plant:\*  
11: SP\_Rodent:\*  
12: SP\_Virus:\*  
13: SP\_Vertebrate:\*  
14: SP\_Unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Hit No.	Score	Query Match	Length	DB ID	Description
1	943	51.0	192	2	Q9S3P3
2	927	50.1	191	2	F70818
3	921	49.8	192	2	Q9R7B1
4	908	49.1	193	2	Q9R8S3
5	893	48.3	200	2	Q9R0R9
6	885	47.8	182	2	Q9R7B2
7	876	47.4	193	2	P94234
8	876	47.4	194	2	Q31122
9	876	47.4	201	2	Q9R0R8
10	863	46.6	191	2	Q44726
11	853	46.1	177	2	Q44979
12	843.5	45.6	178	2	Q9REH7
13	841	45.5	195	2	Q31123
14	829	44.8	175	2	Q9R7B4
15	826	44.6	178	2	Q44995
16	823	44.5	178	2	Q44979
17	802	43.4	163	2	Q9R7A9
18	774.5	41.9	211	2	Q44720
19	773.5	41.8	193	2	P94237

20	773	41.8	159	2	Q9R7B0	Q9R7B0 borrelia bu
21	739	39.9	178	2	Q44994	Q44994 borrelia bu
22	709	38.3	210	2	Q44719	Q44719 borrelia bu
23	707	38.2	192	2	Q9S3P2	Q9S3P2 borrelia bu
24	696.5	37.6	211	2	Q44977	Q44977 borrelia bu
25	692.5	37.4	211	2	Q44976	Q44976 borrelia bu
26	690.5	37.3	193	2	P94242	P94242 borrelia bu
27	686.5	37.1	211	2	Q44976	Q44976 borrelia bu
28	686	37.1	190	2	P94244	P94244 borrelia bu
29	686	37.1	202	2	Q9K1M5	Q9K1M5 borrelia af
30	685.5	37.1	219	2	Q44671	Q44671 borrelia af
31	683.5	36.9	191	2	P94245	P94245 borrelia bu
32	683	36.9	193	2	Q31115	Q31115 borrelia bu
33	681	36.8	212	2	Q44705	Q44705 borrelia bu
34	675	36.5	194	2	P94247	P94247 borrelia bu
35	670	36.2	212	2	Q44727	Q44727 borrelia bu
36	667	36.1	201	2	P96571	P96571 borrelia ja
37	666.5	36.0	193	2	P94231	P94231 borrelia bu
38	666.5	36.0	211	2	Q926C7	Q926C7 borrelia va
39	666	36.0	201	2	P96573	P96573 borrelia ja
40	664.5	35.9	193	2	P94233	P94233 borrelia bu
41	664.5	35.9	202	2	P96508	P96508 borrelia af
42	664.5	35.9	209	2	Q44883	Q44883 borrelia bu
43	664.5	35.9	211	2	Q49577	Q49577 borrelia af
44	664	35.9	203	2	O50624	O50624 borrelia af
45	664	35.9	212	2	Q44670	Q44670 borrelia af

## ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
Q9S3P3	Q9S3P3	Q9S3P3			
AC	Q9S3P3	Q9S3P3			
DT	01-MAY-2000	(Tremblrel. 13, Created)			
DT	01-MAY-2000	(Tremblrel. 13, Last sequence update)			
DT	01-JUN-2001	(Tremblrel. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_Taxid-139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-IP2;				
RX	MEDLINE=96296448; PubMed=8709845;				
RA	Lavey I., Gibbs C.P., Schuster R., Dorner F.;				
RT	"Evidence for lateral transfer and recombination in OspC variation in Lyme disease Borrelia."				
RL	MOL. Microbiol. 18:257-269(1995).				
DR	EMBL: L42887; AAB36995.1; "				
DR	InterPro: IPR001800; Lipoprotein_6.				
DR	Pfam: PF01441; Lipoprotein_6; 1.				
DR	ProDom: PD001149; Lipoprotein_6; 1.				
FT	NON_TER				
FT	TER				
SO	SEQUENCE				
QY	Query Match				
QY	Best Local Similarity				
QY	Matches				
QY	3	MACNSGKDGNTSANSADSEVKGNLTETIKKTTDSNAVLAVKEVEALLSSIDEIAKAIG 62			
QY	1	CNNSGKDGNTSANSADSEVKGNLTETIKKTTDSNAVLAVKEVEALLSSIDEIAKAIG 60			
QY	63	KKHONNGLDENNNHNSGLAGAYASTPLIKOKLDGLKNEGLKEKIDAKKCSFTTNKL 122			
QY	61	KKHONNGLDENNNHNSGLAGAYASTPLIKOKLDGLKNEGLKEKIDAKKCSFTTNKL 120.			
QY	123	KEKHTDLGKEGVTDADAKAEALIKANGTKTGAEELGKLFESVEVLSKAKEMLANSVKEL 182			

```

Db 121 KEKHTDGLGEGYTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVKEL 180
OY 183 TSPVVAESPKKP 194
Db 181 TSPVVAESPKKP 192

RESULT 2
P70818
ID P70818 PRELIMINARY: PRT: 191 AA.
AC P70818:
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN (FRAGMENT).
OSpc.
OC Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RN
RP SEQUENCE FROM N.A.
RC STRAIN-2-1498 CA4;
RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: 181131; AAB06569.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match 50.1%; Score 927; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 1.4e-36;
Matches 186; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 NNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAKIG 63
ID 1 NNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAKIG 60
Db 1 NNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAKIG 60
OY 64 KIHQNNGLDTEENNHNHNSLAGAVAI STLKOKLDGLKNGELKEKIDAAKCKSEFTFNK 123
Db 61 KIHQNNGLDTEENNHNHNSLAGAVAI STLKOKLDGLKNGELKEKIDAAKCKSEFTFNK 120
OY 124 EKHTDGLGEGYTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVKEL 183
121 EKHTDGLGEGYTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVKEL 180
OY 184 SPVVAESPKKP 194
Db 181 SPVVAESPKKP 191

RESULT 3
O9R7B1
ID O9R7B1 PRELIMINARY: PRT: 192 AA.
AC O9R7B1:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OC Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RN
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA MEDLINE-97478003; PubMed-9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RA "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: 091798; AAB81895.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AAE764 CRC64;

Query Match 49.8%; Score 921; DB 2; Length 192;
Best Local Similarity 97.9%; Pred. No. 2.7e-36;
Matches 186; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 3 ISCNNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAK 62
OY 61 IGKTIHQNGLDTEENNHNHNSLAGAVAI STLKOKLDGLKNGELKEKIDAAKCKSEFTFN 120
Db 63 IGKTIHQNGLDTEENNHNHNSLAGAVAI STLKOKLDGLKNGELKEKIDAAKCKSEFTFN 122
OY 121 KKEKHTDGLGEGYTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVK 180
Db 123 KKEKHTDGLGEGYTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVK 182
OY 181 ELTSPVVAES 190
Db 183 ELTSPVVAES 192

RESULT 4
O9R53
ID O9R53 PRELIMINARY: PRT: 193 AA.
AC O9R53:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OSpc.
OC Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RN
RP SEQUENCE FROM N.A.
RC STRAIN-OC1;
RA MEDLINE-99091544; PubMed-9872945;
RA Wang I.N., Dykhuizen D.E., Oju W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of ospC in a local population of Borrelia
RL burgdorferi sensu stricto.";
RL Genetics 151:15-30(1999).
DR EMBL: AF029860; AAB86543.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20502 MW; 5EFD5AF8986D1E CRC64;

Query Match 49.1%; Score 908; DB 2; Length 193;
Best Local Similarity 97.9%; Pred. No. 1.1e-35;
Matches 183; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 7 ISCNNSGKDGNTSANSADSVKGNLTETNKTTDSNAVLAVKEVEALLSSIDEIAAK 66
OY 61 IGKTIHQNGLDTEENNHNHNSLAGAVAI STLKOKLDGLKNGELKEKIDAAKCKSEFTFN 120

```

```

DB 67 IGGKIHONNGLDTEENNHNHNSLAGAVALSTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 126
OY 121 KLEKEHTDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVK 180
DB 127 KLEKEHTDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVK 186
OY 181 ELTSPV 187
DB 187 ELTSPV 193

```

RESULT 5  
 ID 09R09 PRELIMINARY: PRT: 200 AA.

```

DT 01-MAY-2000 (TREMUREL. 13, Created)
DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
GN OSCP.
  Borrelia burgdorferi (Lyme disease spirochete).
  plasmid cp26.
  Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
  NCBI_TaxID=139.

```

```

RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-267.
RX MEDLINE-20002545; PubMed-10531219;
  Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
  "Population dynamics of a naturally occurring heterogeneous mixture of
  Borrelia burgdorferi clones."
  Infect. Immun. 67:5709-5716(1999).
RL EMBL; AF074464; AAD23911.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BB6B6B CRC64;

```

Query Match 48.3%; Score 893; DB 2; Length 200;  
 Best Local Similarity 97.8%; Pred. No. 5.7e-35;  
 Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MACNSGKDGNTSANSDESVKGNLTETNKKTITDSNAVLAVKEVALLSSIDEIATAKA 60
DB 17 ISCNNSGKDGNTSANSDESVKGNLTETNKKTITDSNAVLAVKEVALLSSIDEIATAKA 76
DB 61 IGGKIHONNGLDTEENNHNHNSLAGAVALSTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
DB 77 IGGKIHONNGLDTEENNHNHNSLAGAVALSTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 136
OY 121 KLEKEHTDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVK 180
DB 137 KLEKEHTDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVK 196
OY 181 ELTSPV 184
DB 197 ELTSPV 200

```

RESULT 6  
 ID 09R7B2 PRELIMINARY: PRT: 182 AA.  
 AC 09R7B2;  
 DT 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.

```

OX NCBI_TaxID=139;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RX MEDLINE-97478003; PubMed-9336916;
  Ras N.M., Postic D., Foretz M., Baranton G.;
  "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?"
  Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
  Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; U91797; AAB81894.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

```

Query Match 47.8%; Score 885; DB 2; Length 182;  
 Best Local Similarity 98.9%; Pred. No. 1.2e-34;  
 Matches 180; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 7 GKDNTSANSDESVKGNLTETNKKTITDSNAVLAVKEVALLSSIDEIATAKAIGKKIH 66
DB 1 GKDNTSANSDESVKGNLTETNKKTITDSNAVLAVKEVALLSSIDEIATAKAIGKKIH 60
OY 67 QNNGCLDTEENNHNHNSLAGAVALSTLIKOKLDGLKNEGLKEKIDAKKCEFTTKLEKH 126
DB 61 QNNGCLDTEENNHNHNSLAGAVALSTLIKOKLDGLKNEGLKEKIDAKKCEFTTKLEKH 120
OY 127 TDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVELTSPV 186
DB 121 TDGLKEGVTDADAKAEALITKANGTKTGAEEGLKFESEVYLSKAKEMLANSVELTSPV 180
OY 187 VA 188
DB 181 VA 182

```

RESULT 7  
 ID P94234 PRELIMINARY: PRT: 193 AA.  
 AC P94234;  
 DT 01-MAY-1997 (TREMUREL. 03, Created)  
 DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSCP.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-28354;  
 RX MEDLINE-96296448; PubMed-8709845;  
 RA Lively I., Gibbs C.P., Schuster R., Dorner F.;  
 "Evidence for lateral transfer and recombination in OSCP variation in  
 Lyme disease Borrelia.";  
 Mol. Microbiol. 18:257-269(1995).  
 RL EMBL; I42895; AAB37003.1; -.  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER  
 SQ SEQUENCE 193 AA; 20411 MW; 05B68720F061E2A0 CRC64;

Query Match	47.4%	Score 876	DB 2	Length 201
Best Local Similarity	100.0%	Pred. No. 3.5e-34		
Matches 180; Conservative	0	Mismatches 0; Indels 0; Gaps 0;		
198	NSNGDGNMTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	257		
199	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	258		
200	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	259		
201	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	260		
202	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	261		
203	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	262		
204	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	263		
205	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	264		
206	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	265		
207	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	266		
208	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	267		
209	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	268		
210	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	269		
211	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	270		
212	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	271		
213	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	272		
214	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	273		
215	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	274		
216	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	275		
217	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	276		
218	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	277		
219	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	278		
220	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	279		
221	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	280		
222	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	281		
223	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	282		
224	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	283		
225	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	284		
226	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	285		
227	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	286		
228	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	287		
229	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	288		
230	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	289		
231	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	290		
232	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	291		
233	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	292		
234	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	293		
235	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	294		
236	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	295		
237	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	296		
238	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	297		
239	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	298		
240	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	299		
241	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	300		
242	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	301		
243	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	302		
244	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	303		
245	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	304		
246	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	305		
247	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	306		
248	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	307		
249	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	308		
250	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	309		
251	NSGKDGNTSANSADSVKGPMLTEISKRTESNNVLAVERIEITLLASIDELATKAIGK	310		
252	NSGKDGNTSANSADSV			



QY 319 GEHAOLGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAAKEMLANSVKEL 377  
 DB 121 GEHAOLGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAAKEMLANSVKEL 179

RESULT 11  
 044999 PRELIMINARY: PRT: 177 AA.

AC 044999;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-TXCH;  
 MEDLINE-95286481; Pubmed-7768799;  
 Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,

RT "Evolution of the Borrelia burgdorferi outer surface protein OSpC."  
 Hansen K.;

RL J. Bacteriol. 177:3036-3044(1995).  
 DR EMBL: X84783; CAA59254.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.

FT NON\_TER 177 177  
 FT SEQUENCE 177 AA; 18800 MW; 156671B9614E7A2D CRC64;

Query Match Best Local Similarity 46.1%; Score 853; DB 2; Length 177;  
 Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 NTSANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETAAKAIGKRIHONNG 70  
 DB 1 NTSANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETAAKAIGKRIHONNG 60  
 QY 71 LPTENNHNSSLGAYAVISTLIKOKLDGLKNEGLKEKIDAAKCSFTFTKLEKHTDGL 130  
 DB 61 LPTENNHNSSLGAYAVISTLIKOKLDGLKNEGLKEKIDAAKCSFTFTKLEKHTDGL 120  
 QY 131 KEGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVY 187  
 DB 121 KRGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVY 177

RESULT 12  
 09REH7 PRELIMINARY: PRT: 178 AA.

AC 09REH7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-B31;  
 MEDLINE-94075528; Pubmed-8253951;  
 RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;

RT "Polymorphism in ospC gene of Borrelia burgdorferi and  
 RT immunoreactivity of OSpC protein: implications for taxonomy and for  
 RT use of OSpC protein as a diagnostic antigen.";  
 RL J. Clin. Microbiol. 31:2570-2576(1993).

DR EMBL: X73622; CAA52001.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 178 178  
 FT SEQUENCE 178 AA; 18894 MW; D619A6AC646F14EB CRC64;

Query Match Best Local Similarity 45.6%; Score 843.5; DB 2; Length 178;  
 Matches 174; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 11 NTSANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETAAKAIGKRIHONNG 70  
 DB 1 NTSANSADSVKGNLTETIKKTTDSNAVLAVKEVEALLSIDETAAKAIGKRIHONNG 60  
 QY 71 LPTENNHNSSLGAYAVISTLIKOKLDGLKNEGLKEKIDAAKCSFTFTKLEKHTD-L 129  
 DB 61 LPTENNHNSSLGAYAVISTLIKOKLDGLKNEGLKEKIDAAKCSFTFTKLEKHTD-SF 120  
 QY 130 KEGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVY 187  
 DB 121 KRGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVY 178

RESULT 13  
 031123 PRELIMINARY: PRT: 185 AA.

AC 031123;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS OSpC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-OC13;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF029872; AAB8655.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.

FT NON\_TER 185 185  
 FT SEQUENCE 185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

Query Match Best Local Similarity 45.5%; Score 841; DB 2; Length 185;  
 Matches 172; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 198 NNSGKDGNTSANSADSVKGNLTETIKKTTESNAVLAVKEIETTLASIDELATKAIGK 257  
 DB 10 NNSGKDGNTSANSADSVKGNLTETIKKTTESNAVLAVKEIETTLASIDELATKAIGK 69  
 QY 258 KIOQNGGLAVAGHNGTLLAGAYTISKLTQKLDGLKNSKLEKIKINAKKCSDFTKKL 317  
 DB 70 KIOQNGGLAVAGHNGTLLAGAYTISKLTQKLDGLKNSKLEKIKINAKKCSDFTKKL 129  
 QY 318 EGEHAOLGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAAKEMLANS 373  
 DB 130 EGEHAOLGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAAKEMLANS 185

RESULT 14  
 09R7B4 PRELIMINARY: PRT: 175 AA.  
 AC 09R7B4;

DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 DE Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TETS;  
 RX MEDLINE=97478003; PubMed=9336916;  
 RA Ras N.M., Postic D., Foretz M., Baranton G.;  
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
 U.S.A.'";  
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TETS;  
 RL Marti-Ras N., Postic D., Foretz M., Baranton G.;  
 Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U91792; AAB81889.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 175 175  
 SQ SEQUENCE 175 AA; 18573 MW; EC059E7BD3AC3250 CRC64;

Query Match 44.8%; Score 829; DB 2; Length 175;  
 Best Local Similarity 97.6%; Pred. No. 4.6e-32;  
 Matches 166; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDONTSANSADSVKPNLTETSKITDSNAVLAVKEVEALSSIDEIAKA 60  
 DB 6 ISCNNSGKDONTSANSADSVKPNLTETSKITDSNAVLAVKEVEALSSIDEIAKA 65  
 QY 61 ICKKIHQNNGLDTENNHNHNSLAGAVAI STLKOKLDGLKNEGLKEIDAKKCSFTFN 120  
 DB 66 ICKKIHQNNGLDTENNHNHNSLAGAVAI STLKOKLDGLKNEGLKEIDAKKCSFTFN 125  
 QY 121 KLEKHTDGLGEGVTDADAKKAEILKANGTKTGAEELGKLFESVLSKA 170  
 DB 126 KLEKHTDGLGEGVTDADAKKAEILKANGTKTGAEELGKLFESVLSKA 175

DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MUL;  
 RX MEDLINE=95286481; PubMed=7766799;  
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
 Hansen K.;  
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC";  
 RL J. Bacteriol. 177:3036-3044(1995).  
 DR EMBL; X84779; CA59250.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 178 178  
 SQ SEQUENCE 178 AA; 18911 MW; EE50CE48EADFLCA CRC64;

Query Match 44.6%; Score 826; DB 2; Length 178;  
 Best Local Similarity 98.8%; Pred. No. 6.5e-32;  
 Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 205 NTSANSADSVKPNLTETSKITESNAVLAVKEIEFLASIDEIATKAIGKIQONG 264  
 DB 1 NTSANSADSVKPNLTETSKITESNAVLAVKEIEFLASIDEIATKAIGKIQONG 60  
 QY 265 LAVEAGHNGTSLAGAVTTISKLITOKLDGKNSKEKKEIENAKKSEDFTKLGEHAOL 324  
 DB 61 LAVEAGHNGTSLAGAVTTISKLITOKLDGKNSKEKKEIENAKKSEDFTKLGEHAOL 120  
 QY 325 GIENVTDENAKKAILITDAKDKGAEELEKLFKAVENLAKAKEMLANSVKEL 377  
 DB 121 GIENVTDENAKKAILITDAKDKGAEELEKLFKAVENLAKAKEMLANSVKEL 173

Search completed: March 18, 2002, 10:10:54  
 Job time: 977 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:31 ; Search time 118.14 Seconds

(without alignments)  
232.615 Million cell updates/sec

Title: US-09-596-746A-34

Perfect score: 1009

Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 371

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1804	99.7	370	22	AAB62715
2	1801	99.6	394	22	AAB62730
3	1559.5	86.2	393	22	AAB62716
4	1556.5	86.0	393	22	AAB62731
5	1533.5	84.8	378	22	AAB62712
6	1529	84.5	374	22	AAB62710
7	1526	84.4	398	22	AAB62728
8	1524.5	84.3	401	22	AAB62738
9	1489	82.3	368	22	AAB62718
10	1488.5	82.3	378	22	AAB62725
11	1487.5	82.2	369	22	AAB62714

12	1486	82.1	392	22	AAB62734	Borrelia sp chimera
13	1480.5	81.8	393	22	AAB62732	Borrelia sp chimera
14	1480.5	81.8	401	22	AAB62733	Borrelia sp chimera
15	1289	71.3	377	22	AAB62713	Borrelia sp chimera
16	1284.5	71.0	373	22	AAB62711	Borrelia sp chimera
17	1281.5	70.8	397	22	AAB62729	Borrelia sp chimera
18	1280	70.8	400	22	AAB62739	Borrelia sp chimera
19	1244.5	68.8	367	22	AAB62719	Borrelia sp chimera
20	1241.5	68.6	391	22	AAB62735	Borrelia sp chimera
21	1225	67.7	384	22	AAB62726	Borrelia sp chimera
22	1222	67.6	408	22	AAB62737	Borrelia sp chimera
23	1199	66.3	560	22	AAB62724	Borrelia sp chimera
24	1177.5	65.1	368	22	AAB62717	Borrelia sp chimera
25	1169.5	64.6	391	22	AAB62736	Borrelia sp chimera
26	1097	60.6	410	22	AAB62740	Borrelia sp chimera
27	1096	60.6	386	22	AAB62727	Borrelia sp chimera
28	920	50.9	190	22	AAB62701	Burgdorferi ospc
29	903	49.9	193	22	AAB62702	Burgdorferi ospc
30	886	49.0	211	18	AAW41822	Burgdorferi ospc
31	885	48.9	190	22	AAB62707	Burgdorferi ospc
32	884	48.9	193	15	AAB60888	Borrelia 257 anti
33	879	48.6	192	15	AAB60889	Borrelia 297 anti
34	852	47.1	176	15	AAB62775	Borrelia 297 anti
35	837	46.3	211	20	AAW93473	Burgdorferi ospc
36	822	45.4	177	15	AAB62774	Borrelia 257 anti
37	702	38.8	466	16	AAR75739	B31 outer surface
38	701	38.8	587	16	AAR75746	B31 Osp-A/antigen
39	692	38.3	466	16	AAR75740	B31 outer surface
40	664	36.7	193	22	AAB62709	Burgdorferi ospc
41	661.5	36.6	192	22	AAB62721	Burgdorferi ospc
42	659.5	36.5	209	22	AAB62720	Burgdorferi ospc
43	658.5	36.4	192	22	AAB62703	Burgdorferi ospc
44	658.5	36.4	210	16	AAW11935	Outer surface prot
45	656.5	36.3	192	15	AAB60886	Borrelia Ip2 Ospc

#### ALIGNMENTS

RESULT 1  
AAB62715 standard; Protein: 370 AA.  
XX AAB62715:  
AC AAB62715:  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospc protein SPQ ID NO: 34.  
XX  
XX Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
XX Chimeric - Borrelia sp.  
XX Chimeric - Borrelia sp.  
XX  
XX WO200078966-A1.  
XX  
XX 28-DEC-2000.  
XX  
XX 19-JUN-2000; 2000MO-US16915.  
XX  
XX 18-JUN-1999; 99US-0140042.  
XX  
XX (UANY ) UNIV NEW YORK STATE RES FOUND.  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX WPI, 2001-050113/06.  
XX N-PDB; AAF29019.  
XX  
XX Compositions of OspC polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -

XX Claim 43; Page 91-92; 160pp; English.  
 PS  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 CC  
 XX  
 SQ Sequence 370 AA;

Query Match 99.7%; Score 1804; DB 22; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-104;  
 Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIG 61  
 1 acnnskgdntsansadesvkgpmlteiskrttidsnavllavkeveallsidelakaig 60  
 QY 62 KIRKNDGSLDNEANRNSLLAGAVTISTLTQKLSKINGSEGLKEKTAARKCSEESTK 121  
 Db kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 120  
 QY 122 LKDNHAQLGIGQVTDENAKKAILKANAGKDKGYEELEKLSGLESLSKAEMLANSVK 181  
 Db lkdnhaglgigvtdenakkailkanaagkdkgyeeleklsgleslskaemlansvk 180  
 QY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSID 241  
 Db eltspvhgnnskgdntsansadesvkgpmlteiskrttesnavlavkevealltidsid 240  
 QY 242 ELAKAIGKIRKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAETEKAKKC 301  
 Db elakaignkirkndvslndeadhngslisgaylislntitkisaikdsgelekaetekkacc 300  
 QY 302 SEEFATLKGHEHTDLGEGVTDNNAKKAILETNNDKTGADLEKLPESVKNLSKAEM 361  
 Db seefatlkgehtdlgegytdnnakkailletnnktgadelkelpesvknlskaem 360  
 QY 362 LTNSVKELTS 371  
 Db ltnsvkelts 370

RESULT 2

AK2730  
 AAB62730 standard; Protein; 394 AA.

AC AAB62730;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 64.  
 XX  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX

DR WPI; 2001-050113/06.  
 DR N-PDB; AAF29034.  
 XX

PT Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX

PS Claim 43; Page 128-129; 160pp; English.  
 XX

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 CC  
 XX

SQ Sequence 394 AA;

Query Match 99.6%; Score 1801; DB 22; Length 394;  
 Best Local Similarity 99.7%; Pred. No. 2.6e-104;  
 Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIG 61  
 Db acnnskgdntsansadesvkgpmlteiskrttidsnavllavkeveallsidelakaig 60  
 QY 25 scnnskgdntsansadesvkgpmlteiskrttidsnavllavkevealltidsid 84  
 Db kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 144  
 QY 62 KIRKNDGSLDNEANRNSLLAGAVTISTLTQKLSKINGSEGLKEKTAARKCSEESTK 121  
 Db kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 120  
 QY 85 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSID 241  
 Db eltspvhgnnskgdntsansadesvkgpmlteiskrttesnavlavkevealltidsid 240  
 QY 122 LKDNHAQLGIGQVTDENAKKAILKANAGKDKGYEELEKLSGLESLSKAEMLANSVK 181  
 Db lkdnhaglgigvtdenakkailkanaagkdkgyeeleklsgleslskaemlansvk 180  
 QY 145 LKDNHAQLGIGQVTDENAKKAILKANAGKDKGYEELEKLSGLESLSKAEMLANSVK 204  
 Db lkdnhaglgigvtdenakkailkanaagkdkgyeeleklsgleslskaemlansvk 204  
 QY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSID 241  
 Db eltspvhgnnskgdntsansadesvkgpmlteiskrttesnavlavkevealltidsid 240  
 QY 205 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSID 264  
 Db eltspvhgnnskgdntsansadesvkgpmlteiskrttesnavlavkevealltidsid 264  
 QY 242 ELAKAIGKIRKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAETEKAKKC 301  
 Db elakaignkirkndvslndeadhngslisgaylislntitkisaikdsgelekaetekkacc 324  
 QY 302 SEEFATLKGHEHTDLGEGVTDNNAKKAILETNNDKTGADLEKLPESVKNLSKAEM 361  
 Db seefatlkgehtdlgegytdnnakkailletnnktgadelkelpesvknlskaem 360  
 QY 325 SEEFATLKGHEHTDLGEGVTDNNAKKAILETNNDKTGADLEKLPESVKNLSKAEM 384  
 Db seefatlkgehtdlgegytdnnakkailletnnktgadelkelpesvknlskaem 384  
 QY 362 LTNSVKELTS 371  
 Db ltnsvkelts 394

RESULT 3

AAB62716  
 ID AAB62716 standard; Protein; 369 AA.

AC AAB62716;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 36.  
 XX  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX

```

XX (UUNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29020.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 94-95; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
XX chimeric ospC proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 369 AA:
XX
Query Match 86.2%; Score 1559.5; DB 22; Length 369;
Best Local Similarity 87.5%; Pred. No. 2.2e-89;
Matches 323; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
XX
OY 2 ACNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLLAVKEVALLSSIDELAKAIG 61
DB 1 acnsgkdgtansadesvkgpnltetiskttidnsnavllavkeveallssidelakalg 60
OY 62 KRIKNGSLDNEANRNESLAGAYTISTLTQKLSKNGSEGLKEKIAAACKSEESTK 121
DB 61 krlkngsl dneanr nesllagaytistltqklskngseglkekaaackseestk 120
OY 122 LKDNHQAOLGIGVTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANSVK 181
DB 121 lkdnhqal givt denakkaillkanaagdkgveeleklsgleslskaakemlansvk 180
OY 182 ELTSPVHGNSGKDGNTSANSADSVKGNLTETISKTTESNAVLLAVKEVETLTLSID 241
DB 181 eltspv hgnsgkdgtansadesvkgpnltetiskttetnsnavllavketltlsid 240
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTIKKISAIKDSGELKAEIEKAKK 300
DB 241 elakkaigkklkqngslaveaghnngtlagaytisklilqkldglnkseklkeienakk 300
OY 301 CSEEFPAKLGERTDLGKEGVTDNNAKKAILLKTNNDKTGADLEKLFESVKNLSKAKE 360
DB 301 cseefkklgehnaqigienvt denakkaillltdaakdkgaaleklfkavenlakaake 360
OY 361 MLTNSVKEL 369
DB 361 mlansvkel 369
XX
RESULT 4
AAB62731
ID AAB62731 standard; Protein: 393 AA.
XX
XX AAB62731;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 66.
XX
XX Borrelia: ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX
XX Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

```

```

XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29035.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 131; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
XX chimeric ospC proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 393 AA:
XX
Query Match 86.0%; Score 1556.5; DB 22; Length 393;
Best Local Similarity 87.3%; Pred. No. 3.7e-89;
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;
XX
OY 2 ACNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLLAVKEVALLSSIDELAKAIG 61
DB 25 acnsgkdgtansadesvkgpnltetiskttidnsnavllavkeveallssidelakalg 84
OY 62 KRIKNGSLDNEANRNESLAGAYTISTLTQKLSKNGSEGLKEKIAAACKSEESTK 121
DB 61 krlkngsl dneanr nesllagaytistltqklskngseglkekaaackseestk 120
OY 85 KRIKNGSLDNEANRNESLAGAYTISTLTQKLSKNGSEGLKEKIAAACKSEESTK 144
DB 84 krlkngsl dneanr nesllagaytistltqklskngseglkekaaackseestk 144
OY 122 LKDNHQAOLGIGVTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANSVK 181
DB 121 lkdnhqal givt denakkaillkanaagdkgveeleklsgleslskaakemlansvk 180
OY 145 LKDNHQAOLGIGVTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANSVK 204
DB 144 lkdnhqal givt denakkaillkanaagdkgveeleklsgleslskaakemlansvk 204
OY 182 ELTSPVHGNSGKDGNTSANSADSVKGNLTETISKTTESNAVLLAVKEVETLTLSID 241
DB 181 eltspv hgnsgkdgtansadesvkgpnltetiskttetnsnavllavketltlsid 240
OY 205 ELTSPVHGNSGKDGNTSANSADSVKGNLTETISKTTESNAVLLAVKEVETLTLSID 264
DB 204 eltspv hgnsgkdgtansadesvkgpnltetiskttetnsnavllavketltlsid 264
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTIKKISAIKDSGELKAEIEKAKK 300
DB 241 elakkaigkklkqngslaveaghnngtlagaytisklilqkldglnkseklkeienakk 300
OY 301 CSEEFPAKLGERTDLGKEGVTDNNAKKAILLKTNNDKTGADLEKLFESVKNLSKAKE 360
DB 301 cseefkklgehnaqigienvt denakkaillltdaakdkgaaleklfkavenlakaake 360
OY 361 MLTNSVKEL 369
DB 361 mlansvkel 393
XX
RESULT 5
AAB62712
ID AAB62712 standard; Protein: 378 AA.
XX
XX AAB62712;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 28.
XX

```

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 XX Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX WO200078966-A1.  
 PN 28-DEC-2000.  
 PD 19-JUN-2000; 2000WO-US16915.  
 PF 18-JUN-1999; 99US-0140042.  
 PR (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.  
 N-PSDB; AAF29016.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

Claim 43; Page 83-84; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 378 AA:

Query Match 84.8%; Score 153.5; DB 22; Length 378;  
 Best Local Similarity 86.3%; Pred. No. 9.5e-88;  
 Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNNGKRGNTSANSADSVKGNPTEISKRTDSNAVLAVKEVALLSSIDEL-AKA 59  
 DB 1 macnngkrgntsansadesvkgpnlteinkktdsnavlavkeveallssideiaaka 60  
 QY 60 ICKIKINDGSLDNEANRNESLAGAVTITLTQKLSKNGSEGLKEKIAAKKCSSEFT 119  
 DB 61 igkkihngngldtenhngslagayalstlkqkldgk-negikeidaakkcsseft 119  
 QY 120 TKLDNHAOLGIGQVTDENAKKAILKANAAGKDKVEELKLSGLESLSKAAKEMLAN 179  
 DB 120 ntkkehthdlygevtadakeaalkngt-ktkgaeelgklfesevlskaakemlans 178  
 QY 180 VKELTSPVVG-----NNSGKDGNTSANSADSVKGNPTEISKRTDSNAVLAVK 231  
 DB 179 vkeltsprvaesppkpsmwmsngkdgntsansadesvkgpnlteiskkltesnavlavk 238  
 QY 232 EYETLLTIDELAKAIGKRIKINDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGEL 291  
 DB 239 eveltltsidelakagkikindvsl dneadngslisgavlislnlttkksaalkds9el 298  
 QY 292 KAEIKAKKCSSEFTAKLKGERTDLKRGVTDNNAKKAITLTNNDKTKGADLEKLFESV 351  
 DB 299 kaelakakcsseftaklkgertdlkrgvtdnnaakkaailtkndktkgadelkelfesv 358  
 QY 352 KNLKRAKEMLTNSVKELTS 371  
 DB 359 knlskaakemltnsvkelts 378

RESULT 6  
 AAB62710  
 ID AAB62710 standard; Protein: 374 AA.  
 XX

AC AAB62710;  
 XX 03-APR-2001 (first entry)  
 DT  
 XX Borrelia sp chimeric ospC protein SEQ ID NO: 24.  
 DE  
 XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 KW  
 XX Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX WO200078966-A1.  
 PN 28-DEC-2000.  
 PD 19-JUN-2000; 2000WO-US16915.  
 PF 18-JUN-1999; 99US-0140042.  
 PR (UANY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.  
 N-PSDB; AAF29014.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

Claim 43; Page 78-79; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 374 AA:

Query Match 84.5%; Score 1529; DB 22; Length 374;  
 Best Local Similarity 87.0%; Pred. No. 1.8e-87;  
 Matches 327; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGNPTEISKRTDSNAVLAVKEVALLSSIDEL-AKAI 60  
 DB 1 acnnskgdgntsansadesvkgpnlteiskkltdsnavlavkeveallssideiaakai 60  
 QY 61 GKIKINDGSLDNEANRNESLAGAVTITLTQKLSKNGSEGLKEKIAAKKCSSEFT 120  
 DB 61 gkkihngngldteyngngslagayalstlkqkldgk-negikeidaakkcsseftn 119  
 QY 121 KLDNHAOLGIGQVTDENAKKAILKANAAGKDKVEELKLSGLESLSKAAKEMLAN 180  
 DB 120 klkeththdlygevtadakeaalkngt-ktkgaeelgklfesevlskaakemlansv 178  
 QY 181 KELTSPVVG-----NNSGKDGNTSANSADSVKGNPTEISKRTDSNAVLAVKEVET 235  
 DB 179 keltspvvaesppkpsmwmsngkdgntsansadesvkgpnlteiskkltesnavlavket 238  
 QY 236 LUTSIDELAKAIGKRIKINDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELAEI 295  
 DB 239 ltsidelakagkikindvsl dneadngslisgavlislnlttkksaalkds9el 298  
 QY 296 EKAKKCSSEFTAKLKGERTDLKRGVTDNNAKKAITLTNNDKTKGADLEKLFESV 355  
 DB 299 ekakcsseftaklkgertdlkrgvtdnnaakkaailtkndktkgadelkelfesv 358  
 QY 356 KAAKEMLTNSVKELTS 371  
 DB 359 kaakemltnsvkelts 374



```

RESULT      7
AAB62728
ID AAB62728 standard; Protein: 398 AA.
XX
AC AAB62728;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 60.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN MO200078966-A1.
XX
PD 28-DEC-2000.
XX
UN 19-JUN-2000; 2000MO-US16915.
PR 18-JUN-1999; 99US-0140042.
PA (UNYK ) UNITV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29032.
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 123; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SO Sequence 398 AA;

Query Match      84.4%; Score 1526; DB 22; Length 398;
Best Local Similarity 86.7%; Pred. No. 2.9e-87;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4.

2 ACNNGSGKGONTGANSADSVKGPNTLEISKTTTSDNAVLAAYEVALSSIDEL-AKA1 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||
25 scnnsgkgngncsaansadesvkgpnltelskttidnsavllaaveallssidealaak1 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
61 GKIKINDGSLDNEANRNESLGAAYTITLTIOKLKLSNGSGLKEKIAARKCSSEFST 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
85 gklknqngldetcyhnngsllygayaistllkqldgik-neglikeidaakkcssetfn 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
121 KLDNDHAOLGICGVTDENAKKAIIKANANAAGKGYEELEKSGLESLSKAAREMIANAV 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
144 kikehtldlgkegvtdadakeallktlgtt-kckgaeeilgklfesyevlskaakemianav 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
181 KELTPPVYHG-----NSNGKDGTSANSADSVKGPNTLEISKTTESNAVYLAVEVEFT 235
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
203 keltpvvaaespaamynsgkdgnctsaansadesvkgpnltelskttlesnavylaavevet 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
236 LITSIDELAKAIGKKIRKDVSLDNEDADHNGSLISAGAYLISNLTITTKISAIRKSGELKAEI 295
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
263 lltstidelakelagkkrlndvsidneadnngslisgagylsnlttkksiksdsgeikacl 322
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
266 EAAKCKSEEFTHKKLGERTDLGKBGVTTDDNNAKKAIIKTNNDKTGCADELEKLPESYKNIS 355

```

Dd		333	ekakccseefcaklkgenthdgkgyvtddnakkalkcnndktkgadelcklfesvknls	382
OY		356	KAAKEMLTNSVKEKLS 371	
Db		383	kaakemltnsvkels 398	
RESULT 8				
AAB62738	ID	AAB62738 standard; Protein: 401 AA.		
XX	AC	AAB62738;		
XX	DT	03-APR-2001 (first entry)		
XX	DE	Borrelia sp chimeric ospC protein SEQ ID NO: 80.		
XX	KW	Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.		
XX	OS	Chimeric - Borrelia sp.		
XX	PN	Wo200078966-A1.		
XX	PD	28-DEC-2000.		
XX	PF	19-JUN-2000; 2000WO-US16915.		
XX	PR	18-JUN-1999; 99US-0140042.		
XX	PA	(UYNY ) UNIV NEW YORK STATE RES FOUND.		
XX	PA	(BROO-) BROOK BIOTECHNOLOGIES INC.		
PI		Dattwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;		
XX		WPI: 2001-050113/06.		
DR		N-PSDB: AAF29042.		
XX	PT	Compositions of OspC polypeptides from strains of Borrelia which cause		
XX	PT	Lyme disease are used to immunize animals and detect immune responses		
PI		to Lyme disease -		
XX		Claim 43; Page 150-151; 160pp; English.		
XX	PS	The present invention provides compositions comprising ospC proteins and		
CC		chimeric ospC proteins from members of the Borrelia genus. These may be		
CC		Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as		
CC		vaccines against Borrelia infection, which is spread by ticks and leads		
CC		to Lyme disease.		
XX		Sequence 401 AA;		
SQ				
OY	Query Match	84.3%; Score 1524.5; DB 22; Length 401;		
	Best Local Similarity	86.0%; Pred. No. 3.7e-87;		
	Matches 326; Conservative 10; Mismatches 32; Indels 11; Gaps 4;			
Db	2 ACNNGSGKDGNTPANSADSESVKCPNLTEISKTTTDSNAVLAVKEYEALLSTIDEL-AKAI 60			
	:			
25	scnmgsqkgntcsansadesvkgpnltkskitdsnavllavkeyeallss-idelaakai 84			
	:			
OY	61 GKIKNDSDLDNEARNNESLLGAYTITSLTIQKLSLINGSGLCKEKIAAARKCSBEFST 120			
	:    :         :             :			
Db	85 gkklnngingldteyhnngslagayaisellkqkidglk-neglkexidaacksetfn 143			
	:       :         :             :			
OY	121 KIKDNHAOLGICGVTDENAKKAILKANAAKGKGYBELKLSGSLESLSKAAKEMLANSV 180			
	:       :         :             :			
Db	144 ktkcghtidjgkegyvtcadakealiktngf-kckgaeejgklfesveviskaakemlanv 202			
OY	181 KELTSPPVYHG-----NSSGKDDNTSANSADSESVKCPNLTEISKTTTESNAVVLAWE 232			
Db	203 keltsppvaespdkpsmwnsgkdgnctcsansadesvkgpnltkskitlesnaavvlawe 262			

QY 233 VETLTSLDELAKAIGKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAKIDSGELK 292  
 |||||  
 Db 263 vetltsidelakagkktknvslndeadhngslisgayslntltkksaidsgekl 322  
 QY 293 AELEKAKKCESEFTAKLGEHTDLGEGVTDNNAKKAIIKTNDKTGGADELEKLEFESVK 352  
 |||||  
 Db 323 aeletekaksceeflaktehtdlgkegytdnakkalktndktgadelkelfesvk 382  
 QY 353 NLSKAKEMLTNSVKELT 371  
 |||||  
 Db 383 nlskaakemltnsvkelts 401

RESULT 9  
 AAB62718  
 ID AAB62718 standard; Protein: 368 AA.

AAB62718;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 40.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 9905-0140042.

PA (UNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI, 2001-050113/06.

DR N-PSDB; AAF29022.

XX Compositions of ospC polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

XX to Lyme disease.

XX Claim 43; Page 99-100; 160pp; English.

PS The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 368 AA;

QY Query Match 82.3%; Score 1489; DB 22; Length 368;  
 Best Local Similarity 83.8%; Pred. No. 5.3e-85;

Matches 311; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGPVLTETSKRTITDSNVLAVKEVALLSIDELA-KAI 60  
 |||||  
 Db 1 acnmsgkdgntsansadesvkgpvlteistkrtitdsnavlavkevelisidelakai 60  
 QY 61 CKKINDGSLDNEANRNSLAGAYTITLTQKLSKNGSEGLKERTIAAKKSEEST 120  
 |||||  
 Db 61 gkktngng-leangskntslisgaysltdlaeklnlvkee-lkekldtakgscstefn 118  
 QY 121 KIKDHAQDGIQGYTDEAKAKIILKANAGKDKGVLELEKLSGSELSLAAKEMLANSV 180

Db 119 klskshavlgldnltddnagrallkhh-ankdgyaeleklfkavsnlskaqdtlknav 177  
 QY 181 KELTSPVHHNSGKDGNTSANSADSVKGPVLTETSKRTITDSNVLAVKEVALLSIDELA 240  
 |||||  
 Db 178 keltspvhhnsgkdgntsansadesvkgpvlteistkrtitdsnavlavkevelisid 237  
 QY 241 DELAKAIGKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAKIDSGELKAELEKAKK 300  
 |||||  
 Db 238 delakagkrtkndvslndeadhngslisgayslntltkksaidsgeklaeletek 297  
 QY 301 CSEFTAKLGEHTDLGEGVTDNNAKKAIIKTNDKTGGADELEKLEFESVKNSLAKAE 360  
 |||||  
 Db 298 cseeflaktehtdlgkegytdnakkalktndktgadelkelfesvknlskaake 357  
 QY 361 MLTNSVKELT 371  
 |||||  
 Db 358 mltnsvkelts 368

RESULT 10

AAB62725  
 ID AAB62725 standard; Protein: 378 AA.

AAB62725;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 54.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 9905-0140042.

PA (UNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI, 2001-050113/06.

DR N-PSDB; AAF29029.

XX Compositions of ospC polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

XX to Lyme disease.

XX Claim 43; Page 114-115; 160pp; English.

PS The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

QY Query Match 82.3%; Score 1488.5; DB 22; Length 378;  
 Best Local Similarity 83.3%; Pred. No. 5.8e-85;

Matches 309; Conservative 31; Mismatches 28; Indels 3; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPVLTETSKRTITDSNVLAVKEVALLSIDELA-KAI 60  
 |||||  
 Db 1 macnsgkdgntsansadesvkgpvlteistkrtitdsnavlavkevelisidelakai 60

QY 61 GKIKNDGSLDNEANRNESLAGAVTISTLTOKLSKNGSEGLKEKIAAKKCEEPST 120  
 DB 61 gkikndgsldneanrnesllagaytistltqklskngseglkekaakcseest 120  
 QY 121 KLDNHAQGIQGVTDENAKKAILKANAAAGDKGVEELEKLSGSLESLSKAAKEMLANSV 180  
 DB 121 kldnhaqgiqgvtdenakkailkanaagdkgyveeleklsgsleslskaakemlansv 180  
 QY 181 KELTSVPHGNNSGKDGNTSANSADSVKGPMLTEISKKTESNAVVLAVKEVETLTST 240  
 DB 181 keltspvphgnns--gdaastnpdesakgnltvlskktldsnafllavkeveallstsi 238  
 QY 241 DELAKKIKGKIKNDVSLDNEADHNGSLISGAYLISNLITKKSIAIKDSGELKAEIERAKK 300  
 DB 239 delakkgkikndvslldneadhnsgslisgayslslitkksiaikdsgelekaietk 297  
 QY 301 CSEEFATKLGEGHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNTLSKAKE 360  
 DB 298 cseqfttkldshaelgigsvqddnakkailktchgtkdgakeleefkleslskaaga 357  
 QY 361 MLTNSVKELTS 371  
 DB 358 altnsvkeltn 368  
 RESULT 11  
 AAB62714  
 ID AAB62714 standard; Protein: 369 AA.  
 AC AAB62714:  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 32.  
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN MO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000MO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 (UYNY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI: 2001-050113/06.  
 DR N-PSDB; AAF29018.  
 XX  
 PT Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 89; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 369 AA;  
 Query Match 82.2%; Score 1487.5; DB 22; Length 369;

Best Local Similarity 84.4%; Pred. No. 6,5e-85;  
 Matches 313; Conservative 24; Mismatches 31; Indels 3; Gaps 3;  
 QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKKTDSNAVVLAVKEVALLSIDELAKAIG 61  
 DB 1 acnnskgdntsansadesvkgpmlteiskktidsnavllavkeveallssidelakaig 60  
 QY 62 KIKINDGSLDNEANRNESLAGAVTISTLTOKLSKNGSEGLKEKIAAKKCEEPSTK 121  
 DB 61 kikindgsldneanrnesllagaytistltqklskngseglkekaakcseestck 120  
 QY 122 KLDNHAQGIQGVTDENAKKAILKANAAAGDKGVEELEKLSGSLESLSKAAKEMLANSV 181  
 DB 121 kldnhaqgiqgvtdenakkailkanaagdkgyveeleklsgsleslskaakemlansv 180  
 QY 182 ELTSPVPHGNNSGKDGNTSANSADSVKGPMLTEISKKTESNAVVLAVKEVETLTST 241  
 DB 181 eltspvphgnnsrkdgnstnsadesvkgpmlteiskktesnavvlavkevecllaaid 240  
 QY 242 ELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKKSIAIKDSGELKAEIERAKK 300  
 DB 241 elackaigkikgn-gleangskntslsgaysldlaeklnvlykne-elkekldetkq 298  
 QY 301 CSEEFATKLGEGHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNTLSKAKE 360  
 DB 299 csteftmklselhavlgldnldnagralikkhankdgaalekikfavenlskaagd 358  
 QY 361 MLTNSVKELTS 371  
 DB 359 tlknsvkelts 369

RESULT 12  
 AAB62734  
 ID AAB62734 standard; Protein: 392 AA.  
 AC AAB62734:  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 72.  
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN MO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000MO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 (UYNY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI: 2001-050113/06.  
 DR N-PSDB; AAF29038.  
 XX  
 PT Compositions of ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 139-140; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 392 AA;

Query Match  
 Best Local Similarity 82.1%; Score 1486; DB 22; Length 392;  
 Matches 310; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLLAVKEVEALLSSIDELAKAI 60  
 Db 25 scmsgkdgntsansadesvkgpnlteiskrttesnavllavkeveallssidelakai 84  
 QY 61 GKTKNDGSLDNEANRNESILAGAYTITLTQKLSKLNSEGLKRIIAAKKCEEFST 120  
 Db 85 gkktgnng-leangskntslsagayalsdllaeklnvllknee-lkekldakqgstefn 142  
 QY 121 KLKNHQAOLGIGVTDENAKKAILKANAAGKDKVEELEKLSGLESLSKAKEMLANSV 180  
 Db 143 klkeheavlgldntdnagralikhn-andkgyaelekrlfavenllskaadqtlknay 201  
 QY 181 KELTSPVHGNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLLAVKEVEETLTST 240  
 Db 202 keltsplvhgnskgdntsansadesvkgpnlteiskrttesnavllavkeveetltst 261  
 QY 241 DELAKAIGKTKINDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELAKAEIRAKK 300  
 Db 262 delakalqkktkndvsldeadngslisgayllsnlttkisaiikdsgelekaekak 321  
 QY 301 CSEEFYAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNSKAKE 360  
 Db 322 cseefaklgehtdlygevtddnakkailkcnndktkgadeleklfesvknlskaake 381  
 QY 361 MLTNSVKELTS 371  
 Db 382 mltnsvkelts 392

## RESULT 13

AAB62732  
 ID AAB62732 standard; Protein; 393 AA.

XX AAB62732;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospc protein SEQ ID NO: 68.

XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

XX Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

XX N-PSDB; AAF29036.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 133-134; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 393 AA;

Query Match  
 Best Local Similarity 84.1%; Score 1484.5; DB 22; Length 393;  
 Matches 312; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLLAVKEVEALLSSIDELAKAI 61  
 Db 25 scmsgkdgntsansadesvkgpnlteiskrttdsnnavllavkeveallssidelakai 84  
 QY 62 KRIKNDGSLDNEANRNESILAGAYTITLTQKLSKLNSEGLKRIIAAKKCEEFSTK 121  
 Db 85 krikndgsldeanrnesllagaytistltqklsklngseglkekliaakcseefstk 144  
 QY 122 LKDNHQAOLGIGVTDENAKKAILKANAAGKDKVEELEKLSGLESLSKAKEMLANSV 181  
 Db 145 lkdnhaqlgigvtdenakkailkanaagkdkveeleklsgleslskaakemlansv 204  
 QY 182 ELTSPVHGNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLLAVKEVEETLTST 241  
 Db 205 eltsplvhgnsrgkdgntsansadesvkgpnlteiskrttesnavllavkeveetltst 264  
 QY 242 ELA-RAIKTKINDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELAKAEIRAKK 300  
 Db 265 elatkaigktkgn-gleangskntslsagayalsdllaeklnvllkne-elkekldakq 322  
 QY 301 CSEEFYAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNSKAKE 360  
 Db 323 cseefnklygehtdlygevtddnakkailkcnndktkgadeleklfesvknlskaqd 382  
 QY 361 MLTNSVKELTS 371  
 Db 383 mltnsvkelts 393

## RESULT 14

AAB62733  
 ID AAB62733 standard; Protein; 401 AA.

XX AAB62733;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospc protein SEQ ID NO: 70.

XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

XX Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

```
DR      N-PSDB:AAF629037.  
XX  
PT     Compositions of Ospc polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses -  
PR  
XX  
PS     Claim 43; Page 136-137; 160pp; English.
```

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

```
CC  
CC  
CC  
CC  
CC  
CC  
Sequence    401 AA:  
  
Query Match           Best Local Similarity   81.8%; Score 1480.5; DB 22; Length 401;  
Matches 307; Conservative 32; Mismatches 28; Indels 3; Gaps 2
```

ACNNGSGKDGNTSANSADSVKGPNTLSEISKTTDSNAVLAVKEVEALLSIDELAKAIG 61  
:::|||||  
semsgkdgntsansadesvkgpnltelstsktidsnavllavkeveallssidelakaig 84  
25 semsgkdgnatsansadesvkgnltelstsktidsnavllavkeveallssidelakaig 84

KIKINDGSLDNEANRNESILAGAYITSTLTQKLSKLNSGEGLREXIAAAKKCSFEFSRK 121  
|||  
klknsgsldeanrnesilagayitstlqtqksklngsegglrekaakcsfeefstk 144

LKDNIHAOICGVTPDENAKKATLKNNACKDKDVPELEKLSGSLESLSKAAREMLANSVK 181  
145 lkdnihaiglgvtodenakkalkkanaegkdkyveeleklsglsleslskaaremlansvk 204  
182 ELTSPVHGNNGSGKGNTSANSADSVKGPNTLSEISKTTESNAVLAVKEVEETLLTSID 241  
Db 205 eltspvgnmg--gdasatnpdesahgpnlvtlsktidtnafllavkeveallsisd 262

EIAKAI GKIRKNKVSDINFAHDNLISGAYI LSNLTTKISA IKDGGELAEIEAKKC 301  
|||  
eiskai gkkirkndgtldneanfnealilagayselttklavl-nseellkkkkeadkc 321

SSEFPAFKLGERTDLGREGVTDNMAKKATLKTNNDKTGADELERKFESYVNKSRAKEM 361  
|||  
322 ssekftcklkdsnaehsigivqdnamkakltkchgtkdygaeeelfkstesiskaaga 381

OY 362 LTNSYVELTS 371  
|||  
Db 382 ltnsvelts 391

RESULT 15  
SUBB62713

ID AAB62713 standard; Protein: 377 AA.  
AACB62713;

AAB62713;  
DT 03-APR-2001 (first entry)  
xx DE  
xx Borrella sp chimeric ospC protein SEQ ID NO: 30.  
xx KM  
xx Borrella; ospC; Lyme disease; vaccine; chimeric protein; tick.  
xx OS  
xx Chimeric - Borrella sp.  
xx OS  
xx Chimeric - Borrella sp.  
PN WO200078966-A1.  
xx PD  
xx 28-DEC-2000.  
xx PF  
xx 19-JUN-2000; 2000MO-US16915.  
xx PR  
xx 18-JUN-1999; 99US-0140042.  
xx

PA	(UNIV ) UNIV NEW YORK STATE RES FOUND.	
PA	(BROO ) BROOK BIOTECHNOLOGIES INC.	
xx		
PI	Dattwyler RJ, Seino G, Dykhuisen D, Luft BJ, Gomes-Solecki M;	
xx		
DR	WPI: 2001-050113/06.	
DR	N-PSDB: AAF29017.	
xx		
PT	Compositions of Ospc polypeptides from strains of Borrelia which cause	
PT	Lyme disease are used to immunize animals and detect immune responses	
PI	to Lyme disease -	
xx		
PS	Claim 43; Page 86-87; 160pp: English.	
CC		
CC	The present invention provides compositions comprising ospc proteins and	
CC	chimeric ospc proteins from members of the Borrelia genus. These may be	
CC	Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as	
CC	vaccines against Borrelia infection, which is spread by ticks and leads	
CC	to Lyme disease.	
xx		
SQ	Sequence 377 AA;	
	Query Match 71.3%; Score 1289; DB 22: Length 377;	
	Best Local Similarity 74.1%; Pred. No. 1.4e-72;	
	Matches 281; Conservative 30; Mismatches 56; Indels 12; Gaps 5.	
QY	1 MACNNSGKGNFNSANSADSVKGNPNTETSKKITTSNAVLAVKEVALLSIDEL- 1	AKA 59
DB	1 macnnsqkgnfsansadesvqgnptetlnkkltidsnavlavkevalllsidselaaka	60
QY	60 ICKKIKNGSLDNEANRNESLSLAGVATTISTLTQKLSKNGSEGLKEKITAARKCSEFS	119
DB	61 igkkihngmgldtemhngslagayalsltlqkldgk-neglikexidaakkcetft	119
QY	120 TKLKDNHAOLGIGQYTDENAKKAILKANAGDKGYEELFKLSGLSELSKAKEKLANs	179
DB	120 nklkkehclqkgytdadakaailkangt-ktkgaeeiqklfesyevliskaakemlans	178
QY	180 VKELTSPVYHG-----NNSGKDNFTSANSADSVKGNPNTETSKKITTSNAVLAVK	231
DB	179 veltspvyaesppkpsmwmnsgkdgntcsansadesvqgnptetlskkltesnavlavk	238
QY	232 EVELTLLSIDELA-KAIGKIRNDVSLDMADHNGSLGAYLISMLTKKISAINDSGE	290
DB	232 etetllasidelatkaikgkqngslavaeqhngcllagaytlslklitqkldgklnsek	298
QY	291 LKAELIEAKKCSBEFTAKLGEHTDLGKGCYTDNNAKKAAILTNDNKTGADELEKLEFS	350
DB	299 lkekenakkcsedftkllegehaglgienyvdnenakkailltdaakdkgaeeleklfka	358
QY	351 VKNLSKAKEMLTNSYKEL 369	
DB	359 venlkaakaemlansvkl 377	

Search completed: March 18, 2002, 09:54:32  
Job time: 330 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:33 ; Search time 55.5 Seconds  
(without alignments)  
150.427 Million cell updates/sec

Title: US-09-596-746A-34  
Perfect score: 1809  
Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 371

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: Issued Patents AA:\*  
2: /cgn2\_6/ptodata/2/1aa/3a\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5a\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6a\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	38.8	466	US-08-235-836C-107	Sequence 107, App
2	701	38.8	588	US-08-235-836C-122	Sequence 122, App
3	692	38.3	466	US-08-235-836C-110	Sequence 110, App
4	658.5	36.4	209	US-09-196-293-15	Sequence 15, App1
5	658.5	36.4	210	US-08-158-353-3	Sequence 3, App1
6	653.5	36.1	210	US-08-235-836C-30	Sequence 30, App1
7	652	36.0	207	US-08-235-836C-36	Sequence 36, App1
8	651.5	36.0	210	US-08-209-603E-15	Sequence 15, App1
9	597.5	33.0	212	US-08-158-353-2	Sequence 2, App1
10	588	32.5	212	US-08-031-293-2	Sequence 2, App1
11	588	32.5	212	US-07-903-580-2	Sequence 2, App1
12	556	30.7	212	US-08-158-353-4	Sequence 4, App1
13	554	30.6	209	US-08-235-836C-32	Sequence 32, App1
14	554	30.6	212	US-09-196-293-11	Sequence 11, App1
15	554	30.6	212	US-08-209-603E-11	Sequence 11, App1
16	551	30.5	212	US-08-235-836C-34	Sequence 34, App1
17	166.5	9.2	1713	US-08-600-982-24	Sequence 24, App1
18	166.5	9.2	1713	PCT-US94-10261A-24	Sequence 24, App1
19	164.5	9.1	296	US-08-320-161-11	Sequence 11, App1
20	164.5	9.1	296	US-08-137-175A-3	Sequence 3, App1
21	164.5	9.1	296	US-08-479-017-3	Sequence 3, App1
22	164.5	9.1	296	US-08-455-829-11	Sequence 11, App1
23	164.5	9.1	296	US-08-235-836C-22	Sequence 22, App1
24	163.5	9.0	2482	US-08-328-254-6	Sequence 6, App1
25	163.5	9.0	3248	US-08-353-700-1	Sequence 1, App1
26	163.5	9.0	3248	PCT-US95-16216-1	Sequence 1, App1
27	159.5	8.8	281	US-07-941-523-23	Sequence 23, App1

28	159.5	8.8	380	US-08-235-836C-118	Sequence 118, App
29	159.5	8.8	393	US-08-235-836C-114	Sequence 114, App
30	159.5	8.8	441	US-08-235-836C-120	Sequence 120, App
31	159.5	8.8	454	US-08-235-836C-116	Sequence 116, App
32	152.5	8.4	1786	US-08-973-462-8	Sequence 8, App1
33	151.5	8.4	3111	US-08-460-309-4	Sequence 4, App1
34	151.5	8.4	3111	US-08-125-077-4	Sequence 4, App1
35	143.5	7.9	688	US-09-141-047-8	Sequence 8, App1
36	140.5	7.8	1027	US-08-446-137B-2	Sequence 2, App1
37	139.5	7.7	1561	US-08-894-017-23	Sequence 23, App1
38	137	7.6	1147	US-08-470-260-5	Sequence 5, App1
39	137	7.6	1147	US-08-471-491-5	Sequence 5, App1
40	137	7.6	1147	US-08-466-662-5	Sequence 5, App1
41	137	7.6	1566	US-08-687-956A-23	Sequence 23, App1
42	137	7.6	3289	US-08-477-451-2	Sequence 2, App1
43	136	7.5	1588	PCT-US93-07261-11	Sequence 11, App1
44	136	7.5	1663	PCT-US93-07261-16	Sequence 16, App1
45	134	7.4	1196	US-08-144-121-4	Sequence 4, App1

# ALIGNMENTS

```

RESULT 1
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESSES:
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-Apr-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107

```

Query Match 38.8%; Score 702; DB 4; Length 466;  
Best Local Similarity 47.5%; Pred. No. 2.2e-43;  
Matches 173; Conservative 54; Mismatches 119; Indels 18; Gaps 6;  
19 ESVGPNLITEKTTSDSNVLAIVKVEALLSIDELAKIKKIKDSDLDNANNE 78

```

Db 100 EYFKEDGKTLVSKVTSMDKSTEEKFNEKEVESEKITTTRDGTIRLYTGTSKSGSKAK 159
QY 79 SLGAGVITSTLITOKLSKL--NSEGLKEKIAAKKCEEFSTKLDKNIAOIGIGVTD 136
Db 160 EYFKGVLGEGTLTAKETTLVYKEGTVLSKNISK---SGEVSVELNDTSSAATKRTAA 215
QY 137 ENAKKAILKANAGK-----DKGYELEKLSGLESISKAKEMLANSVKELTSPVY 188
Db 216 WNDSTSTLTISADSKKTKDVLFDITGTTIVQVYNTAGTSLGGSASET--KNISELKNALK 273
QY 189 HGNSGKDGNTSANSADSVKGPMLTEISKRTESNAVLAKEVEFTLLTSDIEI-AKA 247
Db 274 MANNSGKDGNTSANSADSVKGPMLTEISKRTESNAVLAKEVEFTLLTSDIEI-AKA 333
QY 248 GKRIKNDVSLDNEADHNGSLISGAYLISNLITPKISAIIDSGELAELEKAKKCEEF 307
Db 334 GKRIHONGNDLTEXNHNHSGILAGRVAISTLIKOKLDGLKNEG-LKEKIDAAKCEFT 392
QY 308 KLGEHTDLGKEGVTDDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVK 367
Db 393 KLAKEHTDLGKEGVTDAKAEILKTNNGTKTGAELGLFESVYLSKAKEMLANSVK 452
QY 368 ELTS 371
Db 453 ELTS 456

```

## RESULT 2

```

US-08-235-836C-122
; Sequence 122, Application US/03235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-122

```

Query Match 38.8%; Score 701; DB 4; Length 588;  
Best Local Similarity 46.3%; Pred. No. 3,4e-43;  
Matches 181; Conservative 56; Mismatches 86; Indels 68; Gaps 15;

```

QY 6 SGKGNSTANSADSVKGPMLTEISKRTIDSNVLAKEVEFTLLTSDIEI-AKA 65
Db 119 SSKVTKKQSGTTEETLTKANLID--SKRLTRSNCTLTYSQI--TDADNATKAV-ELTK 171
QY 66 MDGSLDNEANRNESLGAAYTI-----STLIQOKLSK-----LNGSEGLKEKIAAKK 113
Db 172 NSIILEG-----SLVVGKTYVIRKEGVTYTKRELEKDGKVKVFLNDPAGSKK----- 219
QY 114 CSEEFSTKLDKNIAOIGIGVTDENAKK-----ALKANAGKDGVEELEKL 161
Db 220 -----TGKMDSTSTLTJIA-----DSKTKKDLVFLDTGITVQVYNTAG-----TSLGGS 265
QY 162 SGLESLSRAAKEMLANSVKELTSPVHGNNGSKDGNTSANSADSVKGPMLTEISKRT 221
Db 266 ASEIKMLSE-----LKNALK--GHPM--GNNSGKDGNTSANSADSVKGPMLTEISKRT 316
QY 222 ESNAVLAKEVEFTLLTSDIEI-AKAIGKIKNDVSLDNEADHNGSLISGAYLISNLITK 280
Db 317 DSNVLAKEVEFTLLTSDIEI-AKAIGKIKNDVSLDNEADHNGSLISGAYLISNLITK 376
QY 281 KISAIKDSGELAELEKAKKCEEFSTAKLGEHTDLGKEGVTDDNNAKAILKTNNDKTKG 340
Db 377 KLDGLKNEG-LKEKIDAAKCEFTTKLEKHTDLGKEGVTDAKAEILKTNNGTKTKG 435
QY 341 ADELKLFESVKNLSKAKEMLTNSVKELT 371
Db 436 ADELKLFESVKNLSKAKEMLTNSVKELT 466

```

## RESULT 3

```

US-08-235-836C-110
; Sequence 110, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid

```



TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 38.3%; Score 692; DB 4; Length 466;  
Best Local Similarity 48.3%; Pred. No. 1,1e-42;  
Matches 185; Conservative 41; Mismatches 123; Indels 34; Gaps 11;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDEL-  
AKA 59  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDETIAKA 76  
QY 60 IGRKIKNDGSLDNEANRNESLLAGATITSTLTQKLSKLSGSEGLKRIAAKCKSEFS 119  
DB 77 IGRKIHQNNGLDTEYNNHNSLLAGRYAISTLIKOKLDGLK-NGELKEKIDAAKCKSETFT 135  
QY 120 TKLKDHAOLGIGYTDENKKAIIKANAGKDKGVELEKLSGSELSKAKKEMANS 179  
DB 136 NKTKAKHTDLCKGCVTDADAKKAILTKTNGT-KTKGAEELCKLEPSEVVLKAKKEMANS 194  
QY 180 VKELTSPVYHGNNSGKDGNTS-NSADESVKGPNTLEISKITTESNA-VVLAKEY 233  
DB 195 VKELTSPVYAESPKRKQNVSLDEKNYSVDLPGEKKVLSKKNKDKGYDLIATVDKL 254  
QY 234 ETLTTS-DELAIAIGKIKNDVSLDNEADHNSLIGAVLI-SNLTITKISA 284  
DB 255 ELKGTSDKNNSGVLEGVADKSKVKLTJS-DDLGOTTELVREKDEKTLVSKKVT 309  
QY 285 IKDSGELKAEIEKAKCKSEFTAKLGEHTDLCKEGYTDONAKKAILTKNODTKGADEL 344  
DB 310 -KDKSSTEKEFNKEGEVSEKIIITRADG--TRLEYTGIKSDSGKAEVLEKGYVLEGLTA 366  
QY 345 EKLFESV-----NLSKAAR 359  
DB 367 EKTTLVYKEGTVTLTKNISKSCE 389

RESULT 4  
US-09-196-293-15  
Sequence 15, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Soutscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia burgdorferi  
FILE REFERENCE: 738,001US2  
CURRENT APPLICATION NUMBER: US/09/196,293  
EARLIER FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 209  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-15

Query Match 36.4%; Score 658.5; DB 4; Length 209;

Best Local Similarity 76.7%; Pred. No. 1,1e-40;  
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDEL-  
AKA 59  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDETIAKA 76  
QY 60 IGRKIKNDGSLDNEANRNESLLAGATITSTLTQKLSKLSGSEGLKRIAAKCKSEFS 119  
DB 77 IGRKIHQNNGLDTEYNNHNSLLAGRYAISTLIKOKLDGLK-NGELKEKIDAAKCKSETFT 135  
QY 120 TKLKDHAOLGIGYTDENKKAIIKANAGKDKGVELEKLSGSELSKAKKEMANS 179  
DB 136 NKTKAKHTDLCKGCVTDADAKKAILTKTNGT-KTKGAEELCKLEPSEVVLKAKKEMANS 194  
QY 180 VKELTSPVY 188  
DB 195 VKELTSPVY 203

RESULT 5  
US-08-158-353-3  
Sequence 3, Application US/08158353  
Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-3

Query Match 36.4%; Score 658.5; DB 1; Length 210;  
Best Local Similarity 76.7%; Pred. No. 1,1e-40;  
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDEL-  
AKA 59  
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKITDTSNAVLLAVKEVEALLSIDETIAKA 76  
QY 60 IGRKIKNDGSLDNEANRNESLLAGATITSTLTQKLSKLSGSEGLKRIAAKCKSEFS 119  
DB 77 IGRKIHQNNGLDTEYNNHNSLLAGRYAISTLIKOKLDGLK-NGELKEKIDAAKCKSETFT 135

Patent No. 6248538

```

? GENERAL INFORMATION:
? APPLICANT: FUCHS, RENATE
? APPLICANT: WILSK, BETTINA
? APPLICANT: PREAC-MURISIC, VERA
? APPLICANT: MOTZ, MANFRED
? APPLICANT: SOUTSCHECK, ERWIN
? TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
? TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
? NUMBER OF SEQUENCES: 15
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
? STREET: 99 PARK AVENUE
? CITY: NEW YORK
? STATE: NY
? COUNTRY: USA
? ZIP: 10016
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" FLOPPY DISC
? COMPUTER: AT&T - IBM COMPATIBLE
? OPERATING SYSTEM: MS-DOS Version 6.2
? SOFTWARE: ASCII
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/209,603E
? FILING DATE: 10-MAR-1994
? CLASSIFICATION: 436
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/EP90/02282
? FILING DATE: 21-DEC-1990
? APPLICATION NUMBER: US 07/862,535
? FILING DATE: 19-JUN-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: ROBINSON, WILLIAM R.
? REGISTRATION NUMBER: 27,224
? REFERENCE/DOCKET NUMBER: LKR-9217-A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 697-3355
? TELEFAX: (212) 697-5635
? INFORMATION FOR SEQ ID NO: 15:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 210
? TYPE: AMINO ACID
? TOPOLOGY: LINEAR
? MOLECULE TYPE: PROTEIN
? DESCRIPTION: N/A
? HYPOTHETICAL: N/A
? ANTI-SENSE: N/A
? FRAGMENT TYPE: N/A
? ORIGINAL SOURCE:
? ORGANISM: B. BURGDORFERI
? IMMEDIATE SOURCE:
? LIBRARY: DSM 5662
? POSITION IN GENOME: N/A
? FEATURE:
? IDENTIFICATION METHOD: amino acid analysis
? PUBLICATION INFORMATION: N/A
? US-08-209-603E-15

Query Match          36.0%; Score 651.5; DB 4; Length 210;
Best Local Similarity 76.2%; Pred. No. 3.4e-40;
Matches 144; Conservative 11; Mismatches 31; Indels 3; Gaps 3;

```

```

OY 180 VKELTSPV 188
DB 195 VKELTSPV 203

RESULT 9
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
? GENERAL INFORMATION:
? APPLICANT: Padula, Steven J.
? TITLE OF INVENTION: Methods for diagnosing Early Lyme
? TITLE OF INVENTION: Disease
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
? STREET: Two Millitia Drive
? CITY: Lexington
? STATE: MA
? COUNTRY: USA
? ZIP: 02173
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/158,353
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Carroll, Alice O.
? REGISTRATION NUMBER: 33,542
? REFERENCE/DOCKET NUMBER: UCT93-05
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-861-6240
? TELEFAX: 617-861-9540
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 212 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-158-353-2

Query Match          33.0%; Score 597.5; DB 1; Length 212;
Best Local Similarity 69.1%; Pred. No. 2.8e-36;
Matches 134; Conservative 20; Mismatches 35; Indels 5; Gaps 4;

```

```

1 GENERAL INFORMATION:
2 APPLICANT: LIVERY, Ian
3 APPLICANT: DORNER, Freidrich
4 TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
5 TITLE OF INVENTION: PREVENTION OF LYME DISEASE
6 NUMBER OF SEQUENCES: 3
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Foley & Lardner
9 STREET: 3000 K Street, N.W., Suite 500
10 CITY: Washington, D.C.
11 COUNTRY: USA
12 ZIP: 20007-5109
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patent Release #1.0, Version #1.25
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/031,295
22 FILING DATE: 19930312
23 CLASSIFICATION: 435
24
25 APPLICATION DATA:
26 APPLICATION NUMBER: US 07/903,580
27 FILING DATE: 25-JUN-1992
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 07/824,161
31 FILING DATE: 22-JAN-1992
32
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US 07/727,245
35 FILING DATE: 11-JUL-1991
36
37 ATTORNEY/AGENT INFORMATION:
38 NAME: BENT, Stephen A.
39 REGISTRATION NUMBER: 29,768
40 REFERENCE/DOCKET NUMBER: 30472/142 IMMU
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE: (202)672-5300
43 TELEFAX: (202)672-5399
44
45 TELEX: 904136
46
47 INFORMATION FOR SEQ ID NO: 2:
48
49 SEQUENCE CHARACTERISTICS:
50 LENGTH: 212 amino acids
51 TYPE: AMINO ACID
52 TOPOLOGY: linear
53
54 MOLECULE TYPE: protein
55
56 US-08-031-295-2
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
9
```

```

1 APPLICANT: LIVERY, Ian
2 APPLICANT: DORNER, Freidrich
3 TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
4 TITLE OF INVENTION: PREVENTION OF LYME DISEASE
5 NUMBER OF SEQUENCES: 3
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: Foley & Lardner
8 STREET: 1800 Diagonal Road, Suite 500
9 CITY: Alexandria
10 STATE: Virginia
11 COUNTRY: USA
12 ZIP: 22313-0299
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patentin Release #1.0, Version #1.25
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/07/903,580
22 FILING DATE: 19920625
23 CLASSIFICATION: 435
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 07/824,161
26 FILING DATE: 22-JAN-1992
27
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 07/727,245
30 FILING DATE: 11-JUL-1991
31 ATTORNEY/AGENT INFORMATION:
32 NAME: BENT, Stephen A.
33 REGISTRATION NUMBER: 29,768
34 REFERENCE/DOCKET NUMBER: 30472/131 IMMU
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (703)836-9300
37 TELEFAX: (703)683-4109
38
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 212 amino acids
42 TYPE: AMINO ACID
43 TOPOLOGY: linear
44
45 MOLECULE TYPE: protein
46
47 US-07-903-580-2
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951

```

NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brock, Smith & Reynolds, P.C.  
 STREET: Two Millitia Drive  
 CITY: Lexington  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/158,353  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Carroll, Alice O.  
 REGISTRATION NUMBER:: 33,542  
 REFERENCE/DOCKET NUMBER: UCT93-05  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-861-6240  
 TELEFAX: 617-861-9540  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 212 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-158-353-4

```

Query March          30.7%; Score 556; DB 1; Length 212;
Best Local Similarity    67.2%; Pred No. 2.8e-33;
Matches 123; Conservative 19; Mismatches 39; Indels 2; Gaps 2

QY      191 NNSGKDGNT-SANSADSVKGPNLTEISKIKTESNAVVLAVEVEVTLSIDELA-KAIG 248
        ||| | : : | ||| ||||| ||||| : ||| ||||| ||||| ||||| |||||
DB      20 NNSGCGGSASINPADESA KGNPLTEISKIKIDSNAFVLAVKEVELTVSIDELAKAIG 79
QY      249 KKIKADVSLDNEADHNGSLISGAYLISNLTYYTKISAIKDSGELKAIEIAKKCSEFTAK 308
        : || : : | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
DB      80 OKIDNNGLALANNONGSILGAVAISTFLITEKLSTKMLBELKTEIARAKKCSEFTAK 139
OY      309 LKGENTDLGEGEVTDNKKATLKTNNDKTRCADELEKLFESVKNSRAKELTNSYKE 368
        || | ||||| : ||| : ||||| : ||| : ||||| : ||| : |||||
DB      140 LKSGHADGKODATDHAKAAITLKTHATTDKCAKEPKDLFESEVGLKAAOVAALTSYKE 199
        |||
        369 LTS 371
        |||
DB      200 LTS 202

RESULT 13
US-08-235-836C-32
Sequence 32, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF SEQUENCES: Borrella Polypeptides and Uses Therefor 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-32

```

```

Query Match          30.6%: Score 554; DB 4; Length 209;
Best Local Similarity 64.6%: Pred. No. 3,8e-33;
Matches 122; Conservative 27; Mismatches 36; Indels 4; Gaps 3

QY      1 MACNNSGKDGMTSANSAPDESQVKGPNLTETSKKITDSTSNNAVLLAVKEVEALLSIDELA-RA 59
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      17 ISCNSSG--GDPASTNPDESAKGNPLTVTSKRTITDSTNAVLAVKEVEALLSIDELANA 74
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      60 IGGKIKNDGSLDNEANRNESLLAGAYTTSTLTITQKLSKINSBGKLEKTAARKCEERS 119
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      75 IGVTHQNNGLLVANAGONGSLAGAYATSTLTITETKLSKLNSEELNKKTEEAKNHSEAF 134
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      120 TLKLNHNLQIGQGVTDENAKKAITKANAAAGKDGVEEELKSGSLSTSKRAKELNLS 179
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      135 NRLKSHQGLGVAATTDHAKETALIKSNPT-KDKGAKALKDSESVESYLAKAQEALANS 193
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      180 VKELTSPVY 188
       11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      194 VKELTNPVY 202

RESULT 14
US-09-196-293-11
: Sequence 11, Application US/09196293
: Patent No. 6183755
: GENERAL INFORMATION:
:   APPLICANT: Fuchs, Renate
:   APPLICANT: Motz, Manfred
:   APPLICANT: Soutscheck, Erwin
:   APPLICANT: Wilske, Bettina
:   APPLICANT: Preac-Murisc, Vera
:   TITLE OF INVENTION: Active proteins from Borrelia
:   TITLE OF INVENTION: burgdorferi
:   FILE REFERENCE: 738.001US2
:   CURRENT APPLICATION NUMBER: US/09/196,293
:   CURRENT FILING DATE: 1998-11-19
:   EARLIER APPLICATION NUMBER: US 08/209,603
:   EARLIER FILING DATE: 1994-03-10
:   EARLIER APPLICATION NUMBER: US 07/862,535
:   EARLIER FILING DATE: 1992-06-19
:   EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
:   EARLIER FILING DATE: 1990-12-21
:   EARLIER APPLICATION NUMBER: DE P39 42 728.5
:   EARLIER FILING DATE: 1989-12-22
:   EARLIER APPLICATION NUMBER: DE P40 18 988.0
:   EARLIER FILING DATE: 1990-06-13
:

```

NUMBER OF SEQ ID NOS: 16  
SOFTWARE: fastseq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 30.6%; Score 554; DB 4; Length 212;  
Best Local Similarity 67.8%; Pred. No. 3.9e-33;  
Matches 124; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

QY 191 NNSGKDG-NTSANSADSVKGPNTLTKSKTTSNAVYLAKEVEYTLISIDELA-KAIG 248  
DB 20 NNSGKVGILTSNPADESAGKPNLTETSKTIDSNFVLAKEVEYTLISIDELAKKAIG 79  
QY 249 KRIKNDVSLDNEADHNGSLISGAYLISNLTJKKISAIKDSGELKAETEKAKKCESEFTAK 308  
DB 80 QKIDNNGLAALNNONGSLAGAYALSTLTETKSLKKNLELTKETELAKKAKKCESEFTAK 139  
QY 309 LKGEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSYKE 368  
DB 140 LKSGHADLGKODATDHAKAAILKTHTATDKAKKEKDLFESYEGLLAKAQAVALTNSYKE 199  
QY 369 LTS 371  
DB 200 LTS 202

RESULT 15  
US-08-209-603E-11

Sequence 11, Application US/08209603E  
Patent No. 6248538

GENERAL INFORMATION:

APPLICANT: FUCHS, RENATE

APPLICANT: WILSKE, BETTINA

APPLICANT: PREC-MURSTIC, VERA

APPLICANT: MOTZ, MANFRED

APPLICANT: SOUTSCHECK, ERMIN

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS

TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY

STREET: 99 PARK AVENUE

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,603E

FILING DATE: 10-MAR-1994

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP90/02282

FILING DATE: 21-DEC-1990

APPLICATION NUMBER: US 07/862,535

FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LKR-9217-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355

TELEFAX: (212) 557-5635

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 212  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
DESCRIPTION: N/A  
HYPOTHETICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORGANISM: B. BURGDORFERI  
IMMEDIATE SOURCE:  
LIBRARY: DSM 5662  
POSITION IN GENOME: N/A  
FEATURE:  
IDENTIFICATION METHOD: amino acid analysis  
PUBLICATION INFORMATION: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: INTERNAL  
ORIGINAL SOURCE:  
US-08-209-603E-11

Query Match 30.6%; Score 554; DB 4; Length 212;  
Best Local Similarity 67.8%; Pred. No. 3.9e-33;  
Matches 124; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

QY 191 NNSGKDG-NTSANSADSVKGPNTLTKSKTTSNAVYLAKEVEYTLISIDELA-KAIG 248  
DB 20 NNSGKVGILTSNPADESAGKPNLTETSKTIDSNFVLAKEVEYTLISIDELAKKAIG 79  
QY 249 KRIKNDVSLDNEADHNGSLISGAYLISNLTJKKISAIKDSGELKAETEKAKKCESEFTAK 308  
DB 80 QKIDNNGLAALNNONGSLAGAYALSTLTETKSLKKNLELTKETELAKKAKKCESEFTAK 139  
QY 309 LKGEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSYKE 368  
DB 140 LKSGHADLGKODATDHAKAAILKTHTATDKAKKEKDLFESYEGLLAKAQAVALTNSYKE 199  
QY 369 LTS 371  
DB 200 LTS 202

Search completed: March 18, 2002, 09:55:34  
Job time: 347 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:44 ; Search time 621.2 Seconds

(Without alignments)  
165.825 Million cell updates/sec

Title: US-09-596-746A-34

1809

Perfect score: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 371

Sequence:

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending Patents: AA.Maln:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1809	100.0	371	19	US-09-596-746A-34
2	1804	99.7	370	19	US-09-596-746-34
3	1801	99.6	394	19	US-09-596-746-64
4	1801	99.6	395	19	US-09-596-746A-64
5	1564.5	86.5	370	19	US-09-596-746A-36
6	1559.5	86.2	369	19	US-09-596-746-36
7	1556.5	86.0	393	19	US-09-596-746-66
8	1556.5	86.0	394	19	US-09-596-746A-66
9	1534	84.8	375	19	US-09-596-746A-24

10	1533.5	84.8	378	19	US-09-596-746-28	Sequence 28, Appl
11	1533.5	84.8	378	19	US-09-596-746A-28	Sequence 28, Appl
12	1529	84.5	374	19	US-09-596-746-24	Sequence 24, Appl
13	1526	84.4	398	19	US-09-596-746-60	Sequence 60, Appl
14	1524.5	84.3	399	19	US-09-596-746A-60	Sequence 60, Appl
15	1524.5	84.3	401	19	US-09-596-746-80	Sequence 80, Appl
16	1524.5	84.3	402	19	US-09-596-746A-80	Sequence 80, Appl
17	1494	82.6	369	19	US-09-596-746A-40	Sequence 40, Appl
18	1492.5	82.5	370	19	US-09-596-746A-32	Sequence 32, Appl
19	1489	82.3	368	19	US-09-596-746-40	Sequence 40, Appl
20	1488.5	82.3	378	19	US-09-596-746-54	Sequence 54, Appl
21	1488.5	82.3	378	19	US-09-596-746A-54	Sequence 54, Appl
22	1487.5	82.2	369	19	US-09-596-746-32	Sequence 32, Appl
23	1486	82.1	392	19	US-09-596-746-72	Sequence 72, Appl
24	1486	82.1	393	19	US-09-596-746A-72	Sequence 72, Appl
25	1484.5	82.1	393	19	US-09-596-746-68	Sequence 68, Appl
26	1484.5	82.1	394	19	US-09-596-746A-68	Sequence 68, Appl
27	1480.5	81.8	401	19	US-09-596-746-70	Sequence 70, Appl
28	1480.5	81.8	402	19	US-09-596-746A-70	Sequence 70, Appl
29	1289.5	71.3	374	19	US-09-596-746A-26	Sequence 26, Appl
30	1289	71.3	377	19	US-09-596-746-30	Sequence 30, Appl
31	1289	71.3	377	19	US-09-596-746A-30	Sequence 30, Appl
32	1284.5	71.0	373	19	US-09-596-746-26	Sequence 26, Appl
33	1281.5	70.8	397	19	US-09-596-746-62	Sequence 62, Appl
34	1281.5	70.8	398	19	US-09-596-746A-62	Sequence 62, Appl
35	1280	70.8	400	19	US-09-596-746-82	Sequence 82, Appl
36	1280	70.8	401	19	US-09-596-746A-82	Sequence 82, Appl
37	1249.5	69.1	368	19	US-09-596-746A-42	Sequence 42, Appl
38	1244.5	68.8	367	19	US-09-596-746-42	Sequence 42, Appl
39	1241.5	68.6	391	19	US-09-596-746-74	Sequence 74, Appl
40	1241.5	68.6	392	19	US-09-596-746A-74	Sequence 74, Appl
41	1230	68.0	385	19	US-09-596-746A-56	Sequence 56, Appl
42	1225	67.7	384	19	US-09-596-746-56	Sequence 56, Appl
43	1222	67.6	408	19	US-09-596-746A-78	Sequence 78, Appl
44	1222	67.6	409	19	US-09-596-746A-78	Sequence 78, Appl
45	1199	66.3	560	19	US-09-596-746-52	Sequence 52, Appl

## ALIGNMENTS

RESULT 1  
US-09-596-746A-34  
; Sequence 34, Application US/09596746A  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seimost, Gerald  
; APPLICANT: Dykhuisen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746A  
; CURRENT FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OSCP Chimera  
US-09-596-746A-34

Query Match 100.0%; Score 1809; DB 19; Length 371;  
Best local similarity 100.0%; Pred. No. 1.9e-120;  
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MACNNSGKDGTSANSADESVKPNLTETSKRTTDSNAVLAVKVEALLSLDELAKAI 60

Db 1 MACNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Db 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Qy 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Db 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Qy 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Db 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 241 DELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Db 241 DELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Qy 301 CSEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Db 301 CSEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Qy 361 LTNSVKEELTS 371  
Db 361 LTNSVKEELTS 371

RESULT 2  
US-09-596-746-34  
Sequence 34 Application US/09596746  
GENERAL INFORMATION:  
APPLICANT: Datwyler, Raymond J.  
APPLICANT: Selmoist, Gerald  
APPLICANT: Dykhuizen, Daniel  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
FILE OF INVENTION: Groups of Borrelia burgdorferi and  
CURRENT APPLICATION NUMBER: 2631.1002-001  
PRIOR FILING DATE: 2000-06-16  
SOFTWARE OF SEQ ID NOS: US 60/140,042  
SEQ ID NO 34  
LENGTH: 370  
FASTSEQ for Windows Version 4.0  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: OSCP Chimera  
US-09-596-746-34

Query Match  
Best Local Similarity 99.7%; Score 1804; DB 19; Length 370;  
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 ACNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Db 1 ACNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Db 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Qy 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Db 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Qy 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Db 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 241 DELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Db 241 DELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Qy 301 CSEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Db 301 CSEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Qy 361 LTNSVKEELTS 371  
Db 361 LTNSVKEELTS 371

Page  
Db 181 ELTSPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 242 ELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Db 242 ELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Qy 302 SEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Db 302 SEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Qy 362 LTNSVKEELTS 371  
Db 362 LTNSVKEELTS 371

RESULT 3  
US-09-596-746-64  
Sequence 64 Application US/09596746  
GENERAL INFORMATION:  
APPLICANT: Datwyler, Raymond J.  
APPLICANT: Selmoist, Gerald  
APPLICANT: Dykhuizen, Daniel  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
FILE OF INVENTION: Groups of Borrelia burgdorferi and  
CURRENT APPLICATION NUMBER: 2631.1002-001  
PRIOR FILING DATE: 2000-06-16  
SOFTWARE OF SEQ ID NOS: US 60/140,042  
SEQ ID NO 64  
LENGTH: 394  
FASTSEQ for Windows Version 4.0  
ORGANISM: OSCP Chimera  
US-09-596-746-64

Query Match  
Best Local Similarity 99.6%; Score 1801; DB 19; Length 394;  
Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 ACNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Db 25 ACNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Db 61 GKIKNDGSLDNEANRNESLAGATTISTITLTKOKSLKNGSGKKEKIAAACEEEST 120  
Qy 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Db 121 KIKNDHOLGIGVTDENAKKAILKANAAGKDGVELEKLSGSLESLAKAEKEMANSV 180  
Qy 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Db 181 KELTSPPVHGNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLLAVEVEALLSIDELAKAI 60  
Qy 242 ELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Db 242 ELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKKISAIDSGELAEKEMANSV 180  
Qy 302 SEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Db 302 SEFTAKLKEGHTDLGREGVTDNNAKAILKTNNDKTKGADLEKLFESVKNLSKAEM 360  
Qy 362 LTNSVKEELTS 371  
Db 362 LTNSVKEELTS 371

```

RESULT 4
US-09-596-746a-64
; Sequence 64, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 395
; TYPE: PRT
; ORGANISM: OSCP Chimera
; 09-596-746a-64

Query Match          99.6%; Score 1801; DB 19; Length 395;
Best Local Similarity 99.7%; Pred. No. 7,6e-120;
Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSIDELAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSIDELAKAI 85
OY 62 KRIKNDGSLDNEANRNESLAGAVTTSTLTORLSKLNSEGLKEKITAARKCSEESTK 121
DB 86 KRIKNDGSLDNEANRNESLAGAVTTSTLTORLSKLNSEGLKEKITAARKCSEESTK 145
OY 122 LKDNHQAOLIGQVTDENAKKAILKANAAKDKVVEELEKISGLESISKAEMLANSV 181
DB 146 LKDNHQAOLIGQVTDENAKKAILKANAAKDKVVEELEKISGLESISKAEMLANSV 205
OY 182 ELTSPVPHGNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSID 241
DB 206 ELTSPVPHGNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSID 265
OY 242 ELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAKK 301
DB 266 ELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAKK 325
OY 302 SEEFPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKADELEKLPESYKNSKAEM 361
DB 326 SEEFPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKADELEKLPESYKNSKAEM 385
OY 362 LTNVSVELTS 371
DB 386 LTNVSVELTS 395

```

```

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OSCP Chimera
; US-09-596-746a-36

Query Match          86.5%; Score 1564.5; DB 19; Length 370;
Best Local Similarity 87.6%; Pred. No. 3e-103;
Matches 324; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSIDELAKAI 60
DB 1 MACNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSIDELAKAI 60
OY 61 GKRIKNDGSLDNEANRNESLAGAVTTSTLTORLSKLNSEGLKEKITAARKCSEESTK 120
DB 61 GKRIKNDGSLDNEANRNESLAGAVTTSTLTORLSKLNSEGLKEKITAARKCSEESTK 120
OY 121 KLDNHAOLIGQVTDENAKKAILKANAAKDKVVEELEKISGLESISKAEMLANSV 180
DB 121 KLDNHAOLIGQVTDENAKKAILKANAAKDKVVEELEKISGLESISKAEMLANSV 180
OY 181 KELTSPVPHGNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSID 240
DB 181 KELTSPVPHGNNSGKDGNTSANSADSVKGPNLFEISKRTDTSNAVLLAVEVEALLSID 240
OY 241 DELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAK 299
DB 241 DELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAK 300
OY 300 KCSPEPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKADELEKLPESYKNSKAEM 359
DB 301 KCSPEPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKADELEKLPESYKNSKAEM 360
OY 360 EMLTNVSVEL 369
DB 361 EMLTNVSVEL 370

```

```

RESULT 6
US-09-596-746-36
; Sequence 36, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OSCP Chimera
; US-09-596-746-36

Query Match          86.2%; Score 1559.5; DB 19; Length 369;

```

Best Local Similarity 87.5%; Pred. No. 1, 1e-102;  
Matches 322; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 61
Db 1 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 60
OY 62 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 121
Db 61 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 120
OY 122 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 181
Db 121 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 180
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 241
Db 181 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 240
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAIEKAK 300
Db 241 ELATKAIGKKIQONGGLAVEAGHNGTLACAVTISTLTOKLDGLNSEKLEKIEKAK 300
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELEKLFESVKNLSKAKE 360
Db 301 CSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 360
OY 361 MLTNSVKEL 369
Db 361 MLANSVKEL 369

```

## RESULT 7

```

US-09-596-746-66
; Sequence 66, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 393
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-66

```

Query Match 86.0%; Score 1556.5; DB 19; Length 393;  
Best Local Similarity 87.3%; Pred. No. 2e-102;  
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 84
OY 62 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 121
Db 85 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 144
OY 122 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 181
Db 145 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 204

```

```

OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 241
Db 205 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 264
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAIEKAK 300
Db 265 ELATKAIGKKIQONGGLAVEAGHNGTLACAVTISTLTOKLDGLNSEKLEKIEKAK 324
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELEKLFESVKNLSKAKE 360
Db 325 CSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 384
OY 361 MLTNSVKEL 369
Db 385 MLANSVKEL 393

```

## RESULT 8

```

US-09-596-746A-66
; Sequence 66, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 394
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-66

```

Query Match 86.0%; Score 1556.5; DB 19; Length 394;  
Best Local Similarity 87.3%; Pred. No. 2e-102;  
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 61
Db 26 SCNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSIDELAKAIG 85
OY 62 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 121
Db 86 KIKINDSLDNEANRNESLACAVTISTLTOKLSKNGSEGLKETAARKCSEESTK 145
OY 122 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 181
Db 146 LKDNHQAOLGIGVTDENAKKAILKANAAGDKVVEELEKLSGLESISKAKEMLANSVK 205
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 241
Db 206 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAWEVEALLSID 265
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAIEKAK 300
Db 266 ELATKAIGKKIQONGGLAVEAGHNGTLACAVTISTLTOKLDGLNSEKLEKIEKAK 325
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELEKLFESVKNLSKAKE 360
Db 326 CSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 385
OY 361 MLTNSVKEL 369
Db 386 MLANSVKEL 394

```

RESULT 9  
US-09-596-746a-24  
; Sequence 24, Application US/09596746a  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Selinost, Gerald  
; APPLICANT: Dykhuzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746a  
; CURRENT FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746a-24

Query Match 84.8%; Score 1534; DB 19; Length 375;  
Best Local Similarity 87.0%; Pred. No. 7.5e-101;  
Matches 328; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSIDEL-  
DB 1 MACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSIDEL-  
QY 60 ICKRIKNDGSLDNEANRNSLLAGAYTISTLTOKLSKNGSEGLKRIAAKCSFEFS  
DB 61 ICKRIHONNGLDTEENNNGSLAGAVASTLIKOKLDGLK-NEGLEKIDAAKCSFEFT  
QY 120 TKLKNHQAOLIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLSKAKKEMLANS  
DB 120 NKLAKKHTDLGEGVTDADAKKAILKTNGT-KTKGAELGKLFESYEVLSKAKKEMLANS  
QY 180 VKELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAVKEVE  
DB 179 VKELTSPVYAESPPAVYNNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAVKEVE  
QY 235 TLITSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTKKISAKDGGEL  
DB 239 TLITSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTKKISAKDGGEL  
QY 295 IEKAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDKTKGADELEKLFESYKNL  
DB 299 IEKAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDKTKGADELEKLFESYKNL  
QY 355 SKAKKEMLTNSVKELTS 371  
DB 359 SKAKKEMLTNSVKELTS 375

RESULT 10  
US-09-596-746-28  
; Sequence 28, Application US/09596746  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Selinost, Gerald  
; APPLICANT: Dykhuzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746-28

Query Match 84.8%; Score 1533.5; DB 19; Length 378;  
Best Local Similarity 86.3%; Pred. No. 8.3e-101;  
Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSIDEL-  
DB 1 MACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSIDEL-  
QY 60 ICKRIKNDGSLDNEANRNSLLAGAYTISTLTOKLSKNGSEGLKRIAAKCSFEFS  
DB 61 ICKRIHONNGLDTEENNNGSLAGAVASTLIKOKLDGLK-NEGLEKIDAAKCSFEFT  
QY 120 TKLKNHQAOLIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLSKAKKEMLANS  
DB 120 NKLAKKHTDLGEGVTDADAKKAILKTNGT-KTKGAELGKLFESYEVLSKAKKEMLANS  
QY 180 VKELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAVK  
DB 179 VKELTSPVYAESPPAVYNNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAVK  
QY 232 EVELTSLTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTKKISAKDGGEL  
DB 239 EVELTSLTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTKKISAKDGGEL  
QY 292 KAEIERAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDKTKGADELEKLFESY  
DB 299 KAEIERAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDKTKGADELEKLFESY  
QY 352 KNLAKKEMLTNSVKELTS 371  
DB 359 KNLAKKEMLTNSVKELTS 378

RESULT 11  
US-09-596-746a-28  
; Sequence 28, Application US/09596746a  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Selinost, Gerald  
; APPLICANT: Dykhuzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746a  
; CURRENT FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera

US-09-596-746a-28

Query Match 84.8%; Score 1533.5; DB 19; Length 378;  
Best Local Similarity 86.3%; Pred. No. 8.3e-101;  
Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNNSGCDGNTSANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSIDEL-AKA 59  
DB 1 MACNNSGCDGNTSANSADSVKGPMLTEINKRTDTSNAVLAVKEVEALLSIDELAAKA 60  
QY 60 GKKIKNGSLDNEANRNEISLAGAYTSTLTOKLSKNGSEGLEKIEIAAKCSEFET 119  
DB 61 GKKITHONNGLDTEENHNGSLAGAYTSTLTOKLDGLK-NEGLEKIDAKKCEFTT 119  
QY 120 TKLDNHAOLGIGQYTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANS 179  
DB 120 NKLKEKHTDLGEGYTDADAKKAILKANGT-KTKGAELKLFESVEVLSKAKEMLANS 178  
QY 180 VELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVK 231  
DB 179 VELTSPVVAESPCKRPSMNVNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVK 238  
QY 232 EVELTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGEL 291  
DB 239 EVELTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGEL 298  
QY 292 KAEIEKAKKCEFTTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADELEKLFESV 351  
DB 299 KAEIEKAKKCEFTTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADELEKLFESV 358  
QY 352 KNLKRAKEMLTNSVELTS 371  
DB 359 KNLKRAKEMLTNSVELTS 378

RESULT 12  
US-09-596-746-24  
Sequence 24, Application US/09596746

GENERAL INFORMATION:  
APPLICANT: Dattwyler, Raymond J.  
APPLICANT: Seinoz, Gerald  
APPLICANT: Dykhulzen, Daniel  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 374  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: OSpC Chimera  
US-09-596-746-24

Query Match 84.5%; Score 1529; DB 19; Length 374;  
Best Local Similarity 87.0%; Pred. No. 1.7e-100;  
Matches 327; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGCDGNTSANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSIDEL-AKAI 60  
DB 1 ACNNSGCDGNTSANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSIDELAAKAI 60  
QY 61 GKKIKNDGSLDNEANRNEISLAGAYTSTLTOKLSKNGSEGLEKIEIAAKCSEFET 120  
DB 61 GKKIKNDGSLDNEANRNEISLAGAYTSTLTOKLSKNGSEGLEKIEIAAKCSEFET 120

DB 61 GKKITHONNGLDTEENHNGSLAGAYTSTLTOKLDGLK-NEGLEKIDAKKCEFTT 119  
QY 121 TKLDNHAOLGIGQYTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANSV 180  
DB 120 KLEKHTDLGEGYTDADAKKAILKTNGT-KTKGAELKLFESVEVLSKAKEMLANSV 178  
QY 181 KELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 235  
DB 179 KELTSPVVAESPAMVNNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 238  
QY 236 LTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 295  
DB 239 LTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 298  
QY 296 EKAKKCEFTTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADELEKLFESVKNLS 355  
DB 299 EKAKKCEFTTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADELEKLFESVKNLS 358  
QY 356 KAKEMLTNSVELTS 371  
DB 359 KAKEMLTNSVELTS 374

RESULT 13  
US-09-596-746-60  
Sequence 60, Application US/09596746

GENERAL INFORMATION:  
APPLICANT: Dattwyler, Raymond J.  
APPLICANT: Seinoz, Gerald  
APPLICANT: Dykhulzen, Daniel  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596,746  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 60  
LENGTH: 398  
TYPE: PRT  
ORGANISM: OSpC Chimera  
US-09-596-746-60

Query Match 84.4%; Score 1526; DB 19; Length 398;  
Best Local Similarity 86.7%; Pred. No. 3.1e-100;  
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGCDGNTSANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSIDEL-AKAI 60  
DB 25 SCNNSGCDGNTSANSADSVKGPMLTEISKRTDTSNAVLAVKEVEALLSIDELAAKAI 84  
QY 61 GKKIKNDGSLDNEANRNEISLAGAYTSTLTOKLSKNGSEGLEKIEIAAKCSEFET 120  
DB 85 GKKITHONNGLDTEENHNGSLAGAYTSTLTOKLDGLK-NEGLEKIDAKKCEFTT 143  
QY 121 TKLDNHAOLGIGQYTDENAKKAILKANAGKDKGVEELEKLSGLESISKAKEMLANSV 180  
DB 144 KLEKHTDLGEGYTDADAKKAILKTNGT-KTKGAELKLFESVEVLSKAKEMLANSV 202  
QY 181 KELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 235  
DB 203 KELTSPVVAESPAMVNNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 262  
QY 236 LTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 295  
DB 263 LTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 322  
QY 296 EKAKKCEFTTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADELEKLFESVKNLS 355

```

Db 323 EKAKKCEFTAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNLS 382
QY 356 KAKKEMLTNSVKELTS 371
Db 383 KAKKEMLTNSVKELTS 398

```

RESULT 14

```

US-09-596-746a-60
; Sequence 60, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnoist, Gerald
; APPLICANT: Dykhuitzen, Dantial
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospC Chlmera
US-09-596-746a-60

```

```

Query Match 84.4%; Score 1526; DB 19; Length 399;
Best Local Similarity 86.7%; Pred. No. 3.1e-100;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

```

```

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVEALLSIDEL-AKAI 60
Db 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVEALLSIDEL-AKAI 85
QY 61 GKRIKNDGSLDNEANRNESLAGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEST 120
Db 86 GKRIHONNGLDTEYNHNGSLAGAVTISTLTQKLDGLK-NEGKLEKIDAKKCEFTN 144
QY 121 KLDNHAOLGIGVTDENAKKAILKANAAGKDVGELEKLSGLSLSKRAKEMLANSV 180
Db 145 KLEKHTDGLGEGVTDADAKKAILKTNGT-KTKGAELGLFESVEVLSKRAKEMLANSV 203
QY 181 KELTSPVVHG-----NNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVE 235
Db 204 KELTSPVVHESPAVNNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVE 263
QY 236 LKTSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISATKDSGELKAEI 295
Db 264 LKTSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISATKDSGELKAEI 323
QY 296 EKAKKCEFTAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNLS 355
Db 324 EKAKKCEFTAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYKNLS 383
QY 356 KAKKEMLTNSVKELTS 371
Db 384 KAKKEMLTNSVKELTS 399

```

```

RESULT 15
US-09-596-746-80
; Sequence 80, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnoist, Gerald
; APPLICANT: Dykhuitzen, Dantial

```

```

; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospC Chlmera
US-09-596-746-80

```

```

Query Match 84.3%; Score 1524.5; DB 19; Length 401;
Best Local Similarity 86.0%; Pred. No. 4e-100;
Matches 326; Conservative 10; Mismatches 32; Indels 11; Gaps 4;

```

```

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVEALLSIDEL-AKAI 60
Db 25 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKEVEALLSIDEL-AKAI 84
QY 61 GKRIKNDGSLDNEANRNESLAGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEST 120
Db 85 GKRIHONNGLDTEYNHNGSLAGAVTISTLTQKLDGLK-NEGKLEKIDAKKCEFTN 143
QY 121 KLDNHAOLGIGVTDENAKKAILKANAAGKDVGELEKLSGLSLSKRAKEMLANSV 180
Db 144 KLEKHTDGLGEGVTDADAKKAILKTNGT-KTKGAELGLFESVEVLSKRAKEMLANSV 202
QY 181 KELTSPVVHG-----NNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKE 232
Db 203 KELTSPVVHESPAVNNNSGKDGNTSANSADSVKGPNTLEISKRTDSNNAVLAVKE 262
QY 233 VETLLTSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISATKDSGELK 292
Db 263 VETLLTSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISATKDSGELK 322
QY 293 AEIEKAKKCEFTAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYK 352
Db 323 AEIEKAKKCEFTAKLGEHTDLGEGVTDNNAKAILKTNNDKTKGADLEKLFESYK 382
QY 353 NLSKAKEMLTNSVKELTS 371
Db 383 NLSKAKEMLTNSVKELTS 401

```

Search completed: March 18, 2002, 10:08:45  
Job time: 973 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:21 ; Search time 81.02 Seconds

(without alignments)  
292.523 Million cell updates/sec

Title: US-09-596-746A-34

Perfect score: 1809

Sequence: 1 MACNNSCKDQNTSANSDESDS.....KNLSKAKEMLTNSVKELTS 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Pending Patents, AA, New.\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	49.0	211	6	US-09-974-992-5
2	656	36.3	207	6	US-09-974-992-3
3	556	30.7	212	6	US-09-974-992-7
4	187.5	10.4	1304	6	US-09-708-427-15045
5	187.5	10.4	1313	6	US-09-708-427-15044
6	187	10.3	837	6	US-09-815-242-5883
7	187	10.3	837	7	US-10-072-851-5883
8	187	10.3	875	6	US-09-815-242-13080
9	187	10.3	875	7	US-10-072-851-13080
10	182.5	10.1	1881	7	US-10-032-585-7646
11	182.5	10.1	1881	7	US-10-072-851-15590
12	176.5	9.8	1014	6	US-09-708-427-19883
13	176.5	9.8	1018	6	US-09-708-427-19882
14	176.5	9.8	1269	6	US-09-708-427-19881
15	169.5	9.4	1144	6	US-09-708-427-15046
16	168.5	9.3	730	6	US-09-708-427-3731
17	168.5	9.3	805	6	US-09-708-427-3730
18	168.5	9.3	841	6	US-09-708-427-3729
19	162.5	9.0	1690	6	US-09-614-150-10224
20	162.5	9.0	1690	6	US-09-614-150-10311
21	162.5	9.0	1833	6	US-09-614-150-40215
22	160	8.8	2434	6	US-09-815-242-5835
23	160	8.8	2434	7	US-10-072-851-5835
24	160	8.8	6281	6	US-09-815-242-12996
25	160	8.8	6281	7	US-10-072-851-12996

26	158	8.7	2346	6	US-09-614-150-17349	Sequence 17349, A
27	155.5	8.6	1948	7	US-10-032-585-7611	Sequence 7611, Ap
28	155.5	8.6	1948	7	US-10-072-851-15555	Sequence 15555, A
29	154	8.5	2437	6	US-09-815-242-5834	Sequence 5834, Ap
30	154	8.5	2437	7	US-10-072-851-5834	Sequence 5834, Ap
31	153.5	8.5	2056	6	US-09-614-150-4824	Sequence 4824, Ap
32	152.5	8.4	1786	6	US-09-742-096-3	Sequence 3, App11
33	152	8.4	679	6	US-09-614-150-25386	Sequence 25386, A
34	151.5	8.4	3070	6	US-09-961-403-7	Sequence 7, App11
35	151.5	8.4	3110	6	US-09-561-709B-7	Sequence 12713, A
36	149.5	8.3	2344	6	US-10-072-851-12713	Sequence 12713, A
37	149.5	8.3	2344	7	US-10-072-851-12713	Sequence 6213, Ap
38	149	8.2	1639	6	US-09-614-150-6213	Sequence 23910, A
39	148	8.2	1409	6	US-09-815-242-12610	Sequence 12610, A
40	148	8.2	5795	6	US-10-072-851-12610	Sequence 12610, A
41	148	8.1	722	6	US-09-815-242-12888	Sequence 12888, A
42	147	8.1	722	7	US-10-072-851-12888	Sequence 12888, A
43	147	8.1	722	7	US-09-815-242-5803	Sequence 5803, Ap
44	147	8.1	991	6	US-09-815-242-5803	Sequence 5803, Ap
45	147	8.1	991	7	US-10-072-851-5803	Sequence 5803, Ap

## ALIGNMENTS

RESULT 1  
US-09-974-992-5  
; Sequence 5, Application US/0974992  
; GENERAL INFORMATION:  
; APPLICANT: Mathiesen, Marianne J.  
; APPLICANT: Theisen, Michael  
; APPLICANT: Holm, Arne  
; APPLICANT: Ostergaard, Soren  
; TITLE OF INVENTION: Novel Opc-derived peptide fragments  
; FILE REFERENCE: 459-666p  
; CURRENT APPLICATION NUMBER: US/09/974, 992  
; CURRENT FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: 09/180, 089  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: PCT/DK97/00203  
; PRIOR FILING DATE: 1997-05-02  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-974-992-5

Query Match 49.0%; Score 886; DB 6; Length 211;  
Best Local Similarity 97.3%; Pred. No. 5.9e-54;  
Matches 183; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNNSCKDQNTSANSDESYKGNLTREISKRTDSNAVLAVKEVALLSIDELAKAI 60  
DB 17 ISNNSSCKDQNTSANSDESYKGNLTREISKRTDSNAVLAVKEVALLSIDELAKAI 76  
QY 61 GKRIKNGSLDNEANRRESLLAGAVTSTLTOKLSLNGSEGLKERTIAAKKSEEPST 120  
DB 77 GKRIKNGSLDNEANRRESLLAGAVTSTLTOKLSLNGSEGLKERTIAAKKSEEPST 136  
QY 121 KLDNHAQLDIOGTDENAKKAILKANAGKDCVEEIEKLSLESISRAKEMLVNSV 180  
DB 137 KLDNHAQLDIOGTDENAKKAILKANAGKDCVEEIEKLSLESISRAKEMLVNSV 196  
QY 181 KELTSPV 188  
DB 197 KELTSPV 204  
RESULT 2  
US-09-974-992-3

```

RESULT      3
US-09-974-992-7
Sequence 7, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSCP-derived peptide fragments
FILE REFERENCE: 459-666p
CURRENT APPLICATION NUMBER: US/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii
US-09-974-992-7

```

Query Match 30.7%; Score 556; DB 6; Length 212;  
 Best Local Similarity 67.2%; Pred. No. 2.9e-31;  
 Matches 123; Conservative 19; Mismatches 39; Indels 2; Gaps 2

```

RESULT 4
US-09-708-427-15045
: Sequence 15045, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708.427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15045
: LENGTH: 1304
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..1304
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..1304
: OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045

```

Query Match	10.4%	Score 187.5;	DB 6;	Length 1304;
Best Local Similarity	23.4%;	Pred. No. 7.3e-05;		
Matches 102; Conservative	75;	Mismatches 149;	Indels 109;	Gaps 177;

```

0Y      4 NNSGDGNTSANSDES-----VGPPLTEISKTIIDSNAYLLAVK-----EVEA 48
        | : : : : : |
Db      85 NEKLKEALAAQKRHEEF EYEKRAVELDQAOLEVOKRDYTSNNELSEIRSQHALLDISA 14
        | : : : : : |
0Y     49 LLSLDELAKAIKKIKNDGSILONE--NNEESLAGAYTISTITTOR-----LSKLNG 10
        | : : : : : |
Db    145 LLSTTEEL-----ORVXHELMSTADKANKAISHAEATRKIAEHAEKAEALLSELORLTA 19
        | : : : : : |
0Y     101 SEGJEKERIAAARKCSEEPSTRLKDNOHQAUGJGYTDENAKKAIKANAACKDKGELEE- 15
        | : : : : : |
Db    200 LLGSEKEKEAIE--GNEYISKIASE-----TELLRGLEEKVSILTSESKLKOBELYEOULV 25
        | : : : : : |
0Y     160 --KISSGESLSAAKEMLANSVKELTSPVYHGNNSGDKGTNSANSDEVGPNLTETS 21
        | : : : : : |
Db    253 DLEAKKAESCTNSSVEEKNRKHVELEKEY-----EESNRKSASBSME-----SYM 30
        | : : : : : |
0Y     218 KRTIESNAVY-----LAKEVEETLITSIDE-----LAKAIRKIRINDY- 25
        | : : : : : |
Db    301 KOALAELNHVIHETKSDNAAQKITELEKTTEAOQRDLDEYGRQVIAAEESALENLAVE 36
        | : : : : : |
0Y     256 -----SLDNEDAHNGSLIGAYLINSITFKKISAIRKIDGSELKAELERAKK 30
        | : : : : : |
Db    361 SIKSELEISOEKRBTALONKAATSN-----IQNLDOQRTELSTELERCKVEEENSKK 41
        | : : : : : |
0Y    301 CSREFTAKLGHTLDGEGYTDDNAKAKALKTKTNNDKYGADELEKLESYVNLSSKAKE 36
        | : : : : : |

```

Db 414 DMSLTAL-----QEASTESSEAKATLVQOEELKNC---ESQVDSLKLAKSETNE 462  
 QY 361 ----MTNSVKELTS 371  
 Db 463 KYEKLEDAAREIDS 477

RESULT 5  
 US-09-708-427-15044

; Sequence 15044, Application US/09708427  
 ; GENERAL INFORMATION:  
 ; APPLICANT: N. ALEXANDROV et al.  
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
 ; TITLE OF INVENTION: THEREBY  
 ; FILE REFERENCE: 2750-1243P  
 ; CURRENT APPLICATION NUMBER: US/09/708,427  
 ; CURRENT FILING DATE: 2000-11-09  
 ; NUMBER OF SEQ ID NOS: 85364  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 15044  
 ; LENGTH: 1313  
 ; TYPE: PR  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 1..1313  
 ; OTHER INFORMATION: Xaa is any amino acid  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 1..1313  
 ; OTHER INFORMATION: Ceres Seq. ID 1828627  
 US-09-708-427-15044

Query Match 10.4%; Score 187.5; DB 6; Length 1313;  
 Best Local Similarity 23.4%; Pred. No. 7.3e-05;  
 Matches 102; Conservative 75; Mismatches 149; Indels 109; Gaps 17;

QY 4 NNSGKGNSTANSANADES-----VKPNLTEISKTTSDNAVLAVK-----EVEA 48  
 Db 94 NERKLALAAQKRAESESFEYERFAVELEQAGLEAVQKDVTSKNLESIKSHALDISA 153  
 QY 49 LSSIDELAKAIGKIKINDSLDNEA-NRNEISLAGAYTISTLITOK-----LSKLNG 100  
 Db 154 LSTTEEL-----QRYKHELSMTADAKNKALSHAEKTKIAEIAEAEILASLGRLKA 208  
 QY 101 SEGKERRIAAKKCSSEFTSKLDNHAQIGVTDENAKKAILKANAAGKDGVEELE- 159  
 Db 209 LGSKEKEKEIE--GNEIYVSKLSE-----IELRLGELEKVSILSESLKQEGLEVLKV 261  
 QY 160 --KLSSLESLSRAAKEMLANSVKELTSPVHGNNSGKDGNTSANSADSEYKGNLTETIS 217  
 Db 262 DLEAAKMAESCTNSVVEEMKNVHELEKEV-----EESNRKSSASSESE-----SYM 309  
 QY 218 KRITESNAVY-----LAVKEVELTLTIDE-----LAKAIGKIRKNDV- 255  
 Db 310 KQALAEHLNVHETKSDNAQKEKIELLEKTIIEAQRDLEEGYQVCAKKEASLLENLVE 369  
 QY 256 -----SLDNEADHNSLISGAYLISNLITTKISAIKDSGELAEIERAKK 300  
 Db 370 SIKSELEISOEERTALDNKKAATSN-----IQLNDQRTETLSIELERCKVEEEKSK 422  
 QY 301 CSSEFTAKLKGHTDLGKEVTDNNAKAILKTNNDKTKADELEKLFESYKNSKAAKE 360  
 Db 423 DMSLTAL-----QEASTESSEAKATLVQOEELKNC---ESQVDSLKLAKSETNE 471  
 QY 361 ----MTNSVKELTS 371  
 Db 472 KYEKLEDAAREIDS 486

RESULT 6  
 US-09-815-242-5883  
 ; Sequence 5883, Application US/09815242

; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; TITLE OF INVENTION: Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5883  
 ; LENGTH: 837  
 ; TYPE: PR  
 ; ORGANISM: Staphylococcus aureus  
 US-09-815-242-5883

Query Match 10.38%; Score 187; DB 6; Length 837;  
 Best Local Similarity 22.28%; Pred. No. 4.4e-05;  
 Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

QY 11 NTSANSADSEYKGNLFEISKRTDS-----NAVLLAVEEALSSIDELAKAIGKIK 65  
 Db 278 NTAFTNLKNGIODONTIKOGVFTDADAKRNATYNAVQAEOIIN-----KAGPPTS 331  
 QY 66 NDG-----SLDNEANRNEISLAGAYTISTLITOKLSKLNGS--EGLKEKIAAARKC 114  
 Db 332 KDGVEETALENVQRAKNELNGNQN-VANAQKTYAKNALNLTSINNAQKELASQIEGATTV 390  
 QY 115 S--EEFSTKLKD--RHAQIGI-QVTDENAKKAILKANAAGKDGVEELEKIS-----G 163  
 Db 391 AGVAVQVSTJASELTJANSNLONGINDEAATRAAL-----NGTONLEKAKOHANTAI 442  
 QY 164 SLESLSRAAKEMLANSVKELTSPVHGNNSGKDGNTSA-----NSADESY--KGNLT 214  
 Db 443 GLSHLTNAQKRALQOLVQOSTTVAEAGNBEQKANNVDAAMKLSQIDANNTTQONQYTT 502  
 QY 215 EIS--KRITESNAVY-----LAVKEVELTLTIDE-----LAKAIGKIRKNDV- 255  
 Db 503 DASONKKDAYNNAVTQAQIITDPTPTIINQAGOVSTTKNALNGNENLEAAKQ 562  
 QY 233 VETLITSIDELAKAIGKIKINDVSLDNEADHNSLISGAYLISNLITTKISAINDSGLK 292  
 Db 563 ASQSLGSLDNNNAQKQVTVDOINGAHVTDANDQKQANQNLNNTAMGMLKQALADKDATK 622  
 QY 293 AEI-----EKAKK-----CSSEFTAKLKGHTDLG--KEGVTDNNAKAILKTNNDK 337  
 Db 623 ATVFTPDADQAKQAYNTAYTNAENIISKANGCNATQAEVQDAIKOVAAQALNGNANY 682  
 QY 338 TKGADELEKLFESYKNSKAAKEMLTNSVKELTS 371  
 Db 683 OHARDEATALINSSNDLNOAQKDALQOVQNAVTT 716

## RESULT 7

US-10-072-851-5883

Sequence 5883, Application US/10072851

GENERAL INFORMATION:

APPLICANT: Carr, Grant J.

APPLICANT: Xu, H. Howard

APPLICANT: Foulkes, J. Gordon

APPLICANT: Zamudio, Carlos

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Roemer, Terry

APPLICANT: Jiang, Bo

APPLICANT: Boone, Charles

APPLICANT: Bussey, Howard

APPLICANT: Bussey, Howard

TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits

TITLE OF INVENTION: Proliferation

FILE REFERENCE: ELITRA.028A

CURRENT APPLICATION NUMBER: US/10/072,851

CURRENT FILING DATE: 2002-02-08

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

NUMBER OF SEQ ID NOS: 15811

SOFTWARE: FastSeq for Windows version 4.0

SEQ ID NO 5883

LENGTH: 837

TYPE: PRN

ORGANISM: Staphylococcus aureus

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-10-072-851-5883

US-09-815-242-13080

Sequence 13080, Application US/09815242

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows version 4.0

SEQ ID NO 13080

LENGTH: 875

TYPE: PRN

ORGANISM: Staphylococcus aureus

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

US-09-815-242-13080

Query Match 10.3%; Score 187; DB 7; Length 837;  
Best Local Similarity 22.2%; Pred. No. 4.4e-05;  
Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

11 NTSANSADSVKGPNTLEISKTTDS-----NAVLLAVEVEALLSIDELAKAICKK 65  
278 NTAAMLNLKNGIDQNTIKGVNFTDADAKRNAYTNAVVOAEQILN-----KAQGPNTS 331  
66 NGC-----SLDNEARNESILAGAYTISLITOKLSKNGS--EGIKETIAAKKC 114  
332 KCGVETALENVORAKNELNGNON-VANAKTTAKNALNLTSINNACKKALKSQIEGATTV 390  
115 S--EEFSTKLKD-NHAGLGIQ-GVTDENAKKALKANAAKDKGVELEKLS-----G 163  
391 AGVNOVSTASELNTAMSLQNGINDEAATKAL-----NGTQNLKAKOHANTAID 442  
164 SLESISKAKEMLANSVKELTSPVHGNSGKGNNTSA-----NSADESV--KGPNTL 214  
443 GSHLTNNAQKALKOLVOOSTVAAEQNGEOKANNVDAAMDKLROSIDADNATTKQONNT 502  
215 EIS--KITESNNAV-----LAVKE 232  
503 DASQNKDAVNNNAVNTAAGTIIDQTSPTLDPYVINOAGOVSTTKNALNGENMLEAKKQ 562  
233 VETLTSIDELAKAICKKIKNDVSLDNEADHNGSLISGAYLLSNLITTKISAIKDSGEIK 292  
563 ASQSLGSLDNLNNAQKQOYTDQINGAHYVDEANQIKONQNLNTAMGNLKOATADADATK 622  
293 AEI-----EKAK-----CSEFTAKIKGHTDLG--KEGVTDNNAKALKITNNNDK 337  
623 ATVNETDADQAKOQAVNTAVTNAENIISKANGSNATQAEVEQAIKOYNNAKQALNGNNAV 662  
338 TKGADELEKLFESVKNLSKAKEMLTNSVYELTS 371  
683 QHAKDEATLINSNDLNQAKALNQOYQVANTT 716

RESULT 8

Query Match 10.3%; Score 187; DB 6; Length 875;  
Best Local Similarity 22.2%; Pred. No. 4.6e-05;  
Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

11 NTSANSADSVKGPNTLEISKTTDS-----NAVLLAVEVEALLSIDELAKAICKK 65  
315 NTAAMLNLKNGIDQNTIKGVNFTDADAKRNAYTNAVVOAEQILN-----KAQGPNTS 368  
66 NGC-----SLDNEARNESILAGAYTISLITOKLSKNGS--EGIKETIAAKKC 114  
369 KCGVETALENVORAKNELNGNON-VANAKTTAKNALNLTSINNACKKALKSQIEGATTV 427  
115 S--EEFSTKLKD-NHAGLGIQ-GVTDENAKKALKANAAKDKGVELEKLS-----G 163  
428 AGVNOVSTASELNTAMSLQNGINDEAATKAL-----NGTQNLKAKOHANTAID 479  
164 SLESISKAKEMLANSVKELTSPVHGNSGKGNNTSA-----NSADESV--KGPNTL 214  
480 GSHLTNNAQKALKOLVOOSTVAAEQNGEOKANNVDAAMDKLROSIDADNATTKQONNT 539  
215 EIS--KITESNNAV-----LAVKE 232  
540 DASQNKDAVNNNAVNTAAGTIIDQTSPTLDPYVINOAGOVSTTKNALNGENMLEAKKQ 599  
233 VETLTSIDELAKAICKKIKNDVSLDNEADHNGSLISGAYLLSNLITTKISAIKDSGEIK 292  
600 ASQSLGSLDNLNNAQKQOYTDQINGAHYVDEANQIKONQNLNTAMGNLKOATADADATK 659  
293 AEI-----EKAK-----CSEFTAKIKGHTDLG--KEGVTDNNAKALKITNNNDK 337  
660 ATVNETDADQAKOQAVNTAVTNAENIISKANGSNATQAEVEQAIKOYNNAKQALNGNNAV 719  
338 TKGADELEKLFESVKNLSKAKEMLTNSVYELTS 371



```

: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Roemer, Terry
: APPLICANT: Jiang, Bo
: APPLICANT: Boone, Charles
: APPLICANT: Bussey, Howard
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: FILE REFERENCE: ELITRA.028A
: CURRENT APPLICATION NUMBER: US/10/072,851
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: 60/267,636
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 15590
: LENGTH: 1881
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1881)..(1881)
: OTHER INFORMATION: X-any amino acid
: US-10-072-851-15590

```

```

Query Match          10.1%; Score 182.5; DB 7; Length 1881;
Best Local Similarity 22.6%; Pred. No. 0.00026;
Matches 101; Conservative 75; Mismatches 139; Indels 131; Gaps 20;

QY 26 LTESKRTT---DSNAVLL-AVKEVEALL-----SIDELAKAIGKIKIND 67
DB 1443 LTESKRTT---DSNAVLL-AVKEVEALL-----SIDELAKAIGKIKIND 67
QY 68 GSLNEANRNESLLAG-----AVTISTLTQKISLT-----NGSEGLAEKIAAA 111
DB 1502 --FENKRYOETTSKDELVEEKOKELVTLQTELKDRISSEVERERAMLSSESTVIKEYSDK 1559
QY 112 KCSSEESTKLKDHNAGIIGVTDENAKKAILKANAGKDKGVEBELEKLSGSLESLSKA 171
DB 1560 KCSSEESTKLKDHNAGIIGVTDENAKKAILKANAGKDKGVEBELEKLSGSLESLSKA 171
QY 172 AKEMLAISVKELTSPVYHGNSGDKGTSANSADSVKGPVLTETISKT-----TES 223
DB 1605 QLEKEMQLEKASL-----EKHNTESATISIE--KNNQIKELSETIKSLKTELKTSG 1656
QY 224 NAVLAIVEVETLLTSLIDELAKAIGK-----KIKNDVSLDNEADHNGSLISGAYLISNL 277
DB 1657 DALNOSQEKYKTLTKNSDYESKLEKOLEBEKVKSDLOTADE-KLKGITERETALAKSEL 1715
QY 278 ITRKISAIKSGELKA-----EIERAK-----KCSSEFTAKLIGEHDTLCK-- 318
DB 1716 ETVKNSGLSTSELALTKTVKSLKEKEBELQFLGSKNSKELEDYIOK---HSDISEKL 1771
QY 319 EGYVD-----DNKRAILKTNNDKTKGADELE----- 345
DB 1772 KALTDLEKTEKTPGDDSKKLTLELNDLSTKKELETKTQTSKFNLEBKDKETVKNL 1831
QY 346 KLFESYKNISSAKAKEMLTNSVKELTS 371
DB 1832 KELELLKNDNSGAKKELSEKVSLETS 1857

```

```

RESULT 12
US-09-708-427-19883
: Sequence 19883, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1243P

```

```

: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 19883
: LENGTH: 1014
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..1014
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc-feature
: LOCATION: 1..1014
: OTHER INFORMATION: Ceres Seq. ID 1836847
: US-09-708-427-19883

```

```

Query Match          9.8%; Score 176.5; DB 6; Length 1014;
Best Local Similarity 24.4%; Pred. No. 0.0003;
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;

QY 16 SADESIV--KGPVLTETISKTIDSNV--LLAVKEVEALLSIDEL-----AKAIGKKIK 65
DB 516 SADESIV--KGPVLTETISKTIDSNV--LLAVKEVEALLSIDEL-----AKAIGKKIK 65
QY 576 D---LEKINSYEQULAEASGSSSLKELEQTLG-----RLAASVVD---KIKQE 622
DB 66 NDGSLDNEANRNESLLAGVITSTLTQKISLNGSEGLAEKIAAACKSEESTKLKDN 125
QY 126 HAQIIGVTDENAKKAILKANAGKDKGVEBELEKLSGSLESLSKAKEMLAISVKELT 185
DB 623 FDQOQEKSLSSSESELAEITNNOIKTK-IOLEGLIGSSVEKETATLKRLEAETERNQ 681
QY 186 PVYHGNSGDKGTSANSADSVKGPVLTETISKTIDSNV--LLAVKEVEALLSIDEL 243
DB 682 KETESDVLVEKLTTHENOIEYKK---LAHESGVADTRREVELEDAKSLKLNESTTEEL 728
QY 244 A-----KAIGKIKNDVSLDNEADHNGSLISGAYLISNLIRKISATDSSELA-EI 295
DB 739 GACQGLEKSSGLAEVNLNLELANHGS-----EAVELDTKLSALEAEKEQTANEL 791
QY 296 EKAKKCSSEFTAKL--KGEHTDLKREGVTDNNAKAILKTNNDKTKGADELEKLFESVK- 352
DB 792 EAKKTTIEDLTQTLSTGEXLOIOIEKLRVAAKESVLESH-----FEELEKTLSEVNA 845
QY 353 NLSKAKEMLTNSVK--ELTS 371
DB 846 QLEKEMQLEKASL-----EKHNTESATISIE--KNNQIKELSETIKSLKTELKTSG 1656

```

```

RESULT 13
US-09-708-427-19882
: Sequence 19882, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 19882
: LENGTH: 1018
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..1018
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc-feature
: LOCATION: 1..1018

```

OTHER INFORMATION: Ceres Seq. ID 1836846  
US-09-708-427-19882

Query Match 9.8%; Score 176.5; DB 6; length 1018;  
Best Local Similarity 24.4%; Pred. No. 0.0003;  
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;

OY 16 SADES- -KGNLITEISKITDSNAV--LLAVEVEALLSIDEL-----AKAIGKKIK 65  
DB 520 SADESLEOGRIDEATTKMELEALHOSLSIDSEHRLQAMEEFTSRDSEASSITEKLR 579  
OY 66 NCSLDNEANRNESLAGAVTTISTLTOKLSKNGSEGLKEKIAAKKCEESTKIKDN 125  
DB 580 D--LEGKIKSTEBOLAEKSGSSSLKEKLEQTLG-----RLAAESVNE---KLKOE 636  
OY 126 HAOLGIGYTDENAKKAILKANAAGDKGVEELEKLSLESLSKAKEMIANSEKELTS 185  
DB 627 FDQAGKSLSSSESESLAETNNQKIK-IOLEGLIGSGSVEKETALKRLEAEIERFNQ 685  
OY 186 PVYHGNSGKDGNTSANSADSVKGNLITEISKITESNAVVL--AVKEVETLTSIDEL 243  
DB 686 KETESSDLVEKLTHTHEQIEYK---LAHESGVADTRKVELEDALSKLKNLESTIEEL 742  
OY 244 A-----KAIGKKIKNDVSLDNEADHNGSLISGAVLISLITKIKSAIKDSGELKA-EI 295  
DB 743 GAKCQGLEKSGDLAEVNLKLNLELANHGS-----EANELQTKLSALEKKEQTANEL 795  
OY 296 EKAKKCESEFTAKL--KGEHTDLGKEGVTDNNAKAILKTNNDKTKGADELEKLFESVK- 352  
DB 796 EAKSTIEDTLTKQLTSEGEKLOQIEKLRVAAEKSVLESH-----FELEKLTSEVKA 849  
OY 353 NLSKAKEMLTNSVK--ELTS 371  
DB 850 QLKENVENATAVSKVAELTS 870

RESULT 14  
US-09-708-427-19881  
; Sequence 19881, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19881  
; LENGTH: 1269  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1269  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1269  
; OTHER INFORMATION: Ceres Seq. ID 1836845  
US-09-708-427-19881

Query Match 9.8%; Score 176.5; DB 6; length 1269;  
Best Local Similarity 24.4%; Pred. No. 0.0004;  
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;  
OY 16 SADES- -KGNLITEISKITDSNAV--LLAVEVEALLSIDEL-----AKAIGKKIK 65  
DB 771 SADESLEOGRIDEATTKMELEALHOSLSIDSEHRLQAMEEFTSRDSEASSITEKLR 830  
OY 66 NCSLDNEANRNESLAGAVTTISTLTOKLSKNGSEGLKEKIAAKKCEESTKIKDN 125

DB 831 D--LEGKIKSTEBOLAEKSGSSSLKEKLEQTLG-----RLAAESVNE---KLKOE 877  
OY 126 HAOLGIGYTDENAKKAILKANAAGDKGVEELEKLSLESLSKAKEMIANSEKELTS 185  
DB 878 FDQAGKSLSSSESESLAETNNQKIK-IOLEGLIGSGSVEKETALKRLEAEIERFNQ 936  
OY 186 PVYHGNSGKDGNTSANSADSVKGNLITEISKITESNAVVL--AVKEVETLTSIDEL 243  
DB 937 KETESSDLVEKLTHTHEQIEYK---LAHESGVADTRKVELEDALSKLKNLESTIEEL 993  
OY 244 A-----KAIGKKIKNDVSLDNEADHNGSLISGAVLISLITKIKSAIKDSGELKA-EI 295  
DB 994 GAKCQGLEKSGDLAEVNLKLNLELANHGS-----EANELQTKLSALEKKEQTANEL 1046  
OY 296 EKAKKCESEFTAKL--KGEHTDLGKEGVTDNNAKAILKTNNDKTKGADELEKLFESVK- 352  
DB 1047 EAKSTIEDTLTKQLTSEGEKLOQIEKLRVAAEKSVLESH-----FELEKLTSEVKA 1100  
OY 353 NLSKAKEMLTNSVK--ELTS 371  
DB 1101 QLKENVENATAVSKVAELTS 1121

RESULT 15  
US-09-708-427-15046  
; Sequence 15046, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15046  
; LENGTH: 1144  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1144  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1144  
; OTHER INFORMATION: Ceres Seq. ID 1828629  
US-09-708-427-15046

Query Match 9.4%; Score 169.5; DB 6; length 1144;  
Best Local Similarity 29.3%; Pred. No. 0.0011;  
Matches 110; Conservative 42; Mismatches 154; Indels 69; Gaps 18;

OY 15 NSADESVKGNLITEISKITDSNAVLLAVEVEALLSIDELAKAIGKIKINDGSLDNEA 74  
DB 606 NLVDNVANMONIAEESDRLR-----REV-AYLKIDELSTANGGLADVTNLONIS 656  
OY 75 NRNE-----SLAGAVTTISTL--ITOKLSKT-----NGSEGLKEKIAAKKCEEST 120  
DB 657 EENKELERETITLLKKADELSELNESLYDKASKIQTYVQENBELRENETATLKRIEELSK 716  
OY 121 K--LKDHAOLGIGYTDENAKKAILKANAAGDKGVEELEKLSLESLSKAKEMLA 177  
DB 717 LHEITLSQETRIQI-----SNHEKEELKRETAAYLKRIEELSKV-----QEDLL 760  
OY 178 NSVKELTSPVYHGNSGKDGNTSANSADSVKGNLITEISKITESNAVVL--LAVKE 232  
DB 761 NKENELHGVVEIEDL-RSKDSLAKOKKITEELSNFNASILLKE-NELDAVVCENEELSKQ 818  
OY 233 VETLITSIDELAKAIGKIKNDVSLDNEADHNGSLISGAVL--ISNLTIKKISAIKIDS 288  
DB 819 VSTLTK-IDELSDLKOSLIRKEKLOAAYVENETLKAEALSLORIEELTITLKOTLIDKO 877

OY 289 GELKA-----EIEKAKCS--PEFTAKLGEHTDLCKEG----VTDDNKAAILKTND- 336  
DB 878 NELQGVFHEHELELAKAKASSIKRIDELLHLEQSWLEKESEFORVTOENLE--LKTQDAL 934  
OY 337 KTKGADELEKLFESV 351  
DB 935 AAKTIEELSKLKESTL 949

Search completed: March 18, 2002, 09:58:23  
Job time: 381 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds  
(without alignments)  
410.946 Million cell updates/sec

Title: US-09-596-746a-34

Perfect score: 1809

Sequence: 1 MACNNSGRKDGNTSANSADSE.....KNLSKAKEMLTNSVKELTS 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	884	48.9	193	2	S70280
2	883	48.8	211	2	S69918
3	879	48.6	210	2	I40144
4	862	47.7	203	2	I40108
5	849	46.9	178	2	I40106
6	663.5	36.7	210	2	S69927
7	662	36.6	193	2	S70287
8	658.5	36.4	210	2	S70218
9	656	36.3	200	2	I40122
10	656	36.3	207	2	S69919
11	654	36.2	193	2	S70286
12	653	36.1	207	2	S37727
13	651.5	36.0	193	2	S70276
14	649.5	35.9	212	2	I40143
15	647	35.8	207	2	I40271
16	643	35.5	189	2	S70263
17	640	35.4	191	2	I40153
18	638	35.3	200	2	S54198
19	636.5	35.2	211	2	I40268
20	636	35.2	209	2	S69917
21	635	35.1	207	2	I40276
22	628	34.7	191	2	S70284
23	622	34.4	211	2	I40145
24	621	34.3	193	2	S70279
25	615.5	34.0	194	2	S70277
26	613	33.9	192	2	S70285
27	612	33.8	193	2	S70259
28	612	33.8	212	2	S69922
29	610	33.7	176	2	I40121

30	609.5	33.7	190	2	S70273	outer surface prot
31	608	33.6	210	2	S69923	outer surface prot
32	608	33.6	212	2	I40279	outer surface prot
33	607.5	33.6	177	2	I40129	outer surface prot
34	607	33.6	194	2	S70268	outer surface prot
35	606	33.5	178	2	I40125	outer surface prot
36	605.5	33.5	193	2	S70274	outer surface prot
37	603	33.3	178	2	I40104	outer surface prot
38	602.5	33.3	199	2	S54197	outer surface prot
39	599	33.1	209	2	I40273	outer surface prot
40	597	33.0	191	2	S70278	outer surface prot
41	595.5	32.9	211	2	S69932	outer surface prot
42	592	32.7	209	2	I40281	outer surface prot
43	590.5	32.6	214	2	S69916	outer surface prot
44	589.5	32.6	194	2	S70289	outer surface prot
45	588	32.5	209	2	I40142	outer surface prot

## ALIGNMENTS

RESULT 1  
S70280  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70280  
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A>Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
A:Reference number: S70255; PMID:96296448  
A:Accession: S70280  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:LA2868; NID:9858735; PIDD:AB37011.1; PTD:91695226  
A:Experimental source: strain Z57  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 48.9%; Score 884; DB 2; Length 193;  
Best Local Similarity 98.4%; Pred. No. 2,9e-34;  
Matches 163; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 C N N S G R K D G N T S A N S A D S V K G P N L T E I S K R T T D S N A V L A V K E V E A L I S I D E L A K A I G K 62  
Db 1 C N N S G R K D G N T S A N S A D S V K G P N L T E I S K R T T D S N A V L A V K E V E A L I S I D E L A K A I G K 60  
Oy 63 K I N D G S L D N E A N R N E S L A G A V T I S T L T O K L S K L N G S E G L K E K I A A K K C S E E F S T K L 122  
Db 61 K I N D G S L D E A N R N E S L A G A V T I S T L T O K L S K L N G S E G L K E K I A A K K C S E E F S T K L 120  
Oy 123 K D N N A O L G I G V D E N K K A I L K A N A G K D G V E L E K L S G L S L S K A A K E M L A N S V K E 182  
Db 121 K D N N A O L G I G V D E N K K A I L K A N A G K D G V E L E K L S G L S L S K A A K E M L A N S V K E 180  
Oy 183 L T S P V V 188  
Db 181 L T S P V V 186

RESULT 2  
S69918  
outer surface protein C precursor - Lyme disease spirochete (strain PBre)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
A:Variety: strain PBre  
C>Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69918; S72674; I40103  
R:Jaursis-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A>Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre

A:Reference number: 140047; MUID:95395018  
 A:Accession: S69918  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-211 <JAN>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022  
 A:Experimental source: strain P8e  
 R:Roessler, D.  
 Submitted to the EMBL Data Library, September 1994  
 A:Reference number: S72674  
 A:Accession: S72674  
 A:Molecule type: DNA  
 A:Residues: 1-152, 'E', 154-211 <ROE>  
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 48.8%; Score 883; DB 2; Length 211;  
 Best Local Similarity 96.8%; Pred. No. 3.5e-34;  
 Matches 182; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MACNSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 60  
 Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 76  
 QY 61 GKTKKNGSLDNEANRNESILLAGAVTITLITQKLSKLNSEGLKEKIAAKKCEEFST 120  
 Db 77 GKTKKNGSLDNEANRNESILLAGAVTITLITQKLSKLNSEGLKEKIAAKKCEEFST 136  
 QY 121 KLDNHAQLGIGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 180  
 Db 137 KLDNHAQLGIGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 196  
 QY 181 KELTSPPV 188  
 Db 197 KELTSPPV 204

RESULT 3  
 I40144  
 outer surface protein C precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
 C:Accession: I40144; S70282  
 C:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022  
 A:Reference number: 140143; MUID:95154673  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Accession: I40144  
 A:Molecule type: DNA  
 A:Residues: 1-210 <RES>  
 A:Cross-references: EMBL:U04281; NID:9434663; PIDN:AC43297.1; PID:9434664  
 R:Liwey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
 Mol. Microbiol. 18, 257-269, 1995  
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
 A:Reference number: S70255; MUID:96296448  
 A:Accession: S70282  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 19-210 <LIV>  
 A:Cross-references: EMBL:I42893; NID:9858721; PIDN:AMB37001.1; PID:g1695218  
 A:Experimental source: strain 297  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 48.8%; Score 879; DB 2; Length 210;  
 Best Local Similarity 99.4%; Pred. No. 5.3e-34;  
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 191 NSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 250  
 Db 20 NSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 79  
 QY 251 IKNDVSLDNEADHNGSLISGAVLISNLTITKISAIKSGELAEIEKAKKCEEFATLK 310  
 Db 80 IKNDVSLDNEADHNGSLISGAVLISNLTITKISAIKSGELAEIEKAKKCEEFATLK 139  
 QY 311 GEHTDGLKGGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 370  
 Db 140 GEHTDGLKGGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 199  
 QY 371 S 371  
 Db 200 S 200

RESULT 4  
 I40108  
 outer surface protein C - Lyme disease spirochete (strain DK7) (fragment)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
 C:Accession: I40108; S34176  
 R:Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.  
 J. Clin. Microbiol. 31, 2570-2576, 1993  
 A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os  
 A:Reference number: I40105; MUID:94075528  
 A:Accession: I40108  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-203 <RES>  
 A:Cross-references: EMBL:X73625; NID:9313275; PIDN:CAA52004.1; PID:g313276  
 C:Genetics:  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 47.7%; Score 862; DB 2; Length 203;  
 Best Local Similarity 95.7%; Pred. No. 3.1e-33;  
 Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MACNSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 60  
 Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKKTITDSNAVLAVKEVEALLSIDELAKAI 76  
 QY 61 GKTKKNGSLDNEANRNESILLAGAVTITLITQKLSKLNSEGLKEKIAAKKCEEFST 120  
 Db 77 GKTKKNGSLDNEANRNESILLAGAVTITLITQKLSKLNSEGLKEKIAAKKCEEFST 136  
 QY 121 KLDNHAQLGIGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 180  
 Db 137 KLDNHAQLGIGVTDENAKKAILKANAAGKDKGVELEKLSGLSLSKAKEMIANSV 196  
 QY 181 KELTSPPV 187  
 Db 197 KELTSPPV 203

RESULT 5  
 I40106  
 outer surface protein C - Lyme disease spirochete (strain HUR) (fragment)  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
 C:Accession: I40106; S54185  
 R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
 J. Bacteriol. 177, 3036-3044, 1995  
 A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
 A:Reference number: I40104; MUID:95286481  
 A:Accession: I40106  
 A:Molecule type: DNA  
 A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84765; NID:g793789; PIDN:CAA59236.1; PID:g793790  
 A:Gene: ospC  
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 46.9%; Score 849; DB 2; Length 178;  
 Best Local Similarity 99.4%; Pred. No. 1e-32;

Matches 177; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

11 NTSANSADESVKGNLTETSKITDTSNAVLAVKEVALLSIDELAKAIGKIKINGSL 70  
 |||||  
 Db 1 NTSANSADESVKGNLTETSKITDTSNAVLAVKEVALLSIDELAKAIGKIKINGSL 60

Qy 71 DNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEFSTRKLDNHAOLG 130  
 |||||  
 Db 61 DNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEFSTRKLDNHAOLG 120

Qy 131 ICGVTDEMAKKAILEKNAAGKDKGVEELEKLSGLESLSKAEMKLANSVKELTSPV 188  
 |||||  
 Db 121 ICGVTDEMAKKAILEKNAAGKDKGVEELEKLSGLESLSKAEMKLANSVKELTSPV 178

RESULT 6

outer surface protein C precursor - Lyme disease spirochete (strain PKa)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

A:Variety: strain PKa

C>Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000

C:Accession: S69927; S72669

R:Jauris-Heipke, S.; Lieg, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia

A:Reference number: 140047; MUID:95395018

A:Accession: S69927

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-210 <OAU>

A:Cross-references: EMBL:X69589

A:Experimental source: strain PKa

R:Jauris, S.  
 submitted to the EMBL Data Library, February 1994

A:Reference number: S72669

A:Accession: S72669

A:Molecule type: DNA

A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAM>

A:Cross-references: EMBL:X69589

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.7%; Score 663.5; DB 2; Length 210;  
 Best Local Similarity 77.2%; Pred. No. 3.8e-24;

Matches 146; Conservative 10; Mismatches 30; Indels 3; Gaps 3;

Qy 1 MACNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVALLSIDEL-AGA 59  
 :|||

Db 17 ISCNNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVALLSIDELAKA 76  
 :|||

Qy 60 IGGKIKNGSLDNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEF 119  
 |||||

Db 77 IGGKIKNGSLDNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEF 135  
 |||||

Qy 120 TKLKDHAOLGICGVTDENAKKAILEKNAAGKDKGVEELEKLSGLESLSKAEMK 179  
 |||||

Db 136 NKTKDHTDGLGKGVTDADAKKAILTKNGT-KTKGAELKLEPSEVLSKAEMK 194  
 |||||

RESULT 7

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C:Accession: S70287

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di

A:Reference number: S70255; MUID:96296448

A:Accession: S70287

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-193 <LIV>

A:Cross-references: EMBL:I42895; NID:g858723; PIDN:AB37003.1; PID:g1695220

A:Experimental source: strain 28354

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.6%; Score 662; DB 2; Length 193;  
 Best Local Similarity 75.9%; Pred. No. 4.1e-24;  
 Matches 142; Conservative 17; Mismatches 26; Indels 2; Gaps 2;

Qy 3 CNNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVALLSIDELAKAIG 61  
 |||||

Db 1 CNNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVALLSIDELAKAIG 60  
 |||||

Qy 62 KIKINDOSLDNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEF 121  
 |||||

Db 61 KIKINDOSLDNEANRNESILAGAYTISTLITOKLSKLNSEGLKEKIAAKKCEEF 120  
 |||||

Qy 122 LKDNHAOLGICGVTDENAKKAILEKNAAGKDKGVEELEKLSGLESLSKAEMK 181  
 |||||

Db 121 LEGHAOLGIEVTDENAKKAILEKNAAGKDKGVEELEKLSGLESLSKAEMK 179  
 |||||

Qy 182 ELTSPV 188  
 |||||

Db 180 ELTSPV 186  
 |||||

RESULT 8

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000

C:Accession: G70218; I40269; S37726; S70281

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu

; Bowman, C.; Garland, S.; Fujii, C.; Colton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98055943

A:Accession: G70218

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-210 <RES>

A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AA66329.1; PID:g2689901; TIGR:BB

A:Experimental source: strain B31

R:Fukunaga, M.; Hamase, A.

J. Clin. Microbiol. 33, 2415-2420, 1995

A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu

A:Reference number: I40269; MUID:96025162

A:Accession: I40269

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 <RES>

A:Cross-references: GB:D49497; NID:g707092; PIDN:BA08457.1; PID:g769684

R:Jauris-Heipke, S.; Fuchs, R.; Mctz, R.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.

Med. Microbiol. Immunol. 182, 37-50, 1993

A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (os

Qy 121 KLKDHAGLGQGVYDENAKKAILIKANAAGDKVEEELKSGLESSTSKAKELNLSV 180  
 ||||:||||:|:|||||||: |||:||||: |||||: |||  
 Db 134 KLKSHALGLQSQVDDAKKAILIKTHGT-KDKGAKLEELFKLSLSLSTKNAQAALNTSV 192  
 Qy 181 KELTSPVY 188  
 ||||:||||  
 Db 193 KELTNPVY 200  
 RESULT 10  
 S69919  
 Outer surface protein C precursor - Borrelia garinii (strain P7rob)  
 C:Species: Borrelia garinii  
 A:Variety: strain P7rob  
 C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
 C:Accession: S69919  
 R:Jauris-Heipke, S.; Liegel, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek  
 R. Clin. Microbiol. 33, 1860-1866, 1995

Db	17	ISCANSGDGNISANSDESVKGPNIETSKITDSNVLVKEEALSSIDEIAAKA	76	111	A: experimental source: strain p100
					Ar: Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

0V 120 TKIKDNHAOLCTOGVTDENAKKATI KANADSGKDKVEVEI EKI SCST ECT EVA VEVI ANG 170

QY 180 VKELTPV 188

outer surface protein C - Lyme disease spirochete (strain DK6) (fragment)

R;Theisen, M.; Frederiksen, B.; Iebeh, A.M.; Vuust, J.; Hansen, K.

A/Title: Polymorphism in ospC gene of *Borrelia burgdorferi* and immunoreactivity of ospC

A;status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A;Cross-references: EMBL:X73626; NID:g313277; PIDN:CAA52005.1; PID:g313278  
C:Genetics:

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

Best Local Similarity 75.0%; Pred. No. 8e-24;
Matches 141; Conservative 20; Mismatches 23; Indels 4; Gaps 3;

```

```
Db      17 TSCNNCC - CDEACEMNDEECACVCDUT EMTQCVZTEDCNU DUTATWWTUUAUUCOCCTGTCVAT
```

QY 61 GKKIKNDGSLDNEANRNESLAGAYTISTLITQKLSKLNSEGLKEKIAAAKCKSEEFST 120

D<sub>b</sub> 75 GKKIKNDGTLIDNEANRRNESLIAGAVEISKILITQKSYLVN-SEELKEKIKEANDCSEKFTT 133

A; Variety: strain PTH

R. Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek J. Clin. Microbiol. 33, 1860-1866, 1995

A;Accession: S69919

A; Cross-references: EMBL:X83554; NID:g872027; PIDN:CA58544.1; PID:g872028

C;Genetics:  
A:Gene. Osm

F;19-207/Product: outer surface protein C #status predicted

F;19-207/Product: outer surface protein C #status predicted <MAT>

```

Best Local Similarity 75.06; Pred. NO. 8.3e-24;
Matches 141; Conservative 20; Mismatches 23; Indels 4; Gaps 3;

```

```
db      17 TSCNNSG--GDSASTNPDESAGPNITVTSKKTDSNAEILAKKEVEALLSTIDEISKAT 7A
```

[illegible]

121 KLENDHAGUGV I DENNAKAI KANAAGNDAGVEELENLSGLESLSKAANEMLANSV .180

DB 134 K L K D S H A E L G L Q S V Q D D N A K K A L L K T H G T - K D K G A K E L E E L F K S L E S L S K A Q A A L T N S V 192

Db 193 KeltNPV 200

RESULT 11  
070000

C:\Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
c,species: borreliaburgdorferi ( Lyme disease spirochete)

Mol. Microbiol. 18, 257-269, 1995

A; Status: nucleic acid sequence not shown

A;Cross-references: EMBL:L42897; NID:g858728; PIDN:AAB37006.1; PID:g1695222

A; Experimental source: strain 26815  
C; Genetics:

A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.2%: Score 654; DB 2; Length 193;  
Best Local Similarity 74.5%: Pred. No. 9.4e-24;  
Matches 140; Conservative 19; Mismatches 25; Indels 4; Gaps 3;

OY 3 CANSKGDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKAIG 62  
DB 1 CANSKGDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKAIG 60  
OY 63 KIRKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFTKL 122  
DB 61 KINNG-LDVOQNFNAISLGAHTISKLYTEKLSKLNSEGLKEIDAKKCSDDFTKKL 119  
OY 123 KDNHAQGIQ--VTDENAKKAILKANAAGKDKVEELEKLSGLESLSKAKEMLANSV 180  
DB 120 QSSMAQGVAGGATTDDEAKKAILRTNAI-KDKGADELKLFKVSESLAKAQAQALANSV 178  
181 KELTSPPV 188  
179 NELTSPV 186

RESULT 12

outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-May-2000  
C:Accession: S37727  
R:Jauris-Helpe, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)  
A:Reference number: S37726; MUID:93268136  
A:Accession: S37727  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-207 <JAU>  
A:Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA49305.1; PID:g311394  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.1%: Score 653; DB 2; Length 207;  
Best Local Similarity 74.5%: Pred. No. 1.1e-23;  
Matches 140; Conservative 21; Mismatches 23; Indels 4; Gaps 3;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKAI 60  
DB 17 ISCNNSG-GDSASTNPDESAGKGNLTIVISKITDTSNAFLAVKEVEALLSSIDELSKAI 74  
61 GKRIKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 120  
DB 75 GKRIKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 133  
OY 121 KLDNHAQGIQVTDENAKKAILKANAAGKDKVEELEKLSGLESLSKAKEMLANSV 180  
DB 134 KLDNHAQGIQVTDENAKKAILKILKTHTGT-KDKGADELKLFKVSLESLSKAQAALANSV 192  
OY 181 KELTSPPV 188  
DB 193 KELTSPPV 200

RESULT 13

S70276

outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70276  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea

A:Reference number: S70255; MUID:96296448  
A:Accession: S70276  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:U42890; NID:g858718; PIDN:AAB36998.1; PID:g1695215  
A:Experimental source: strain E61  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.0%: Score 651.5; DB 2; Length 193;  
Best Local Similarity 76.5%: Pred. No. 1.2e-23;  
Matches 140; Conservative 17; Mismatches 23; Indels 3; Gaps 3;

OY 191 NNSKDG-NTSANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKAIG 248  
DB 2 NNSKGDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 61  
OY 249 KIRKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 308  
DB 62 KIRKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 120  
OY 309 LKGEHTDLGKGVYDDNAKKAIIKTNNDKTGADELKLFESVKNLSKAKEMLANSV 368  
DB 121 LKDRHTELGKODANDDAKKAIIKTNNGDKTGADELKLFESVKNLSKAKEMLANSV 180  
OY 369 LTS 371  
DB 181 LTS 183

RESULT 14

I40143

outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40143  
R:Stevenson, B.; Barthold, S.W.  
FEMS Microbiol. Lett. 124, 367-372, 1994  
A:Title: Expression and sequence of outer surface protein C among North American isol  
A:Reference number: I40143; MUID:95154673  
A:Accession: I40143  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-212 <RES>  
A:Cross-references: EMBL:U04280; NID:g2314879; PIDN:AAC45542.1; PID:g434662  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.9%: Score 649.5; DB 2; Length 212;  
Best Local Similarity 74.3%: Pred. No. 1.7e-23;  
Matches 142; Conservative 19; Mismatches 25; Indels 5; Gaps 4;

OY 1 MACNNSGKDGNT-SANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKA 59  
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETSKITDTSNAVLAVKEVEALLSSIDELAKA 76  
OY 60 GKRIKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 119  
DB 77 GKRIKNDGSLDNEANRNESSLGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPFT 135  
OY 120 TKLDNHAQGIQ--VTDENAKKAILKANAAGKDKVEELEKLSGLESLSKAKEMLA 177  
DB 136 KLDNHAQGIQVTDENAKKAILKILKTHTGT-KDKGADELKLFKVSLESLSKAQAALANSV 194  
OY 178 NSVKELTSPV 188  
DB 195 NSVKELTSPV 205

## RESULT 15

I40271

outer surface protein C precursor - *Borrelia garinii*

C;Species: Borrelia garinii

C; Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
C; Accession: T40271

C; Accession: I40271  
B: Eukaryota M: Ham

R; Fukunaga, M.; Hammase, A  
Clin Microbiol 33: 20

J. Clin. Microbiol. 33, 2415-2420, 1995

A;Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato  
A;Reference number: 140269; MUID: 96025162

A:Accession: I40271  
A:Reference number: 140269; MUID:96025162

A;Accession: 1402/1  
A;Status: preliminary

A: Molecule type: DNA  
A: status: preliminary; translated from GB/EMBL/DBO

```
A;Molecule type: DNA
A;Residues: 1-207 <R
```

A;Cross-references: GB:

C; Superfamily: Lyme disease spirochete surface protein C

Experimentally: Time increases exponentially with process.

query Match

35.88; Score 647; DB 2; Length 207;  
72.38; Pred. No. 2.1e-23;

Matches 136; Conservative 27; Mismatches 21

[illegible]

QY 1 MACNNSGKDNTSANSADSVKGPNLTEISKKITDSNAVLAVKEVEALLSSIDELAKAI 6C

Db 17 ISCSNG--GDTASTNPDESAGPNLIEISKITDSNAVLAVKEVEALLSSIDELSKAI 74

QY 61 GKKIKNDGSLDNEANRNESLLAGAYTISTLTQKLSKNGSEGLKEIAAKKCSEEST 120

Db 75 GKKIRNDGTLNNEANRNESLIAGAYEISKLTQKLSVLN-SELKEKIKEAKDCSEKFTT 133

QY 121 K L K D N H A Q L G I Q G V T D E N A K K A I L K A N A A G K D K V E E L E K L S G S L E S L S K A K E M L A N S V 180

Db 134 KLRD SHAELGVQNVQDDNAKRAILKTH-GNKKDKAKELKELSESLLENLAKAAQAASSNSV 1922

QY 181 KELTSPV 188

Db 193 KELTSPV 200

```
Search completed: March 18, 2002, 09:56:48
Job time: 371 sec
```

Job time: 371 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:49 ; Search time 39.62 Seconds  
(without alignments)  
343.328 Million cell updates/sec

Title: US-09-596-746a-34  
Perfect score: 1809  
Sequence: 1 MACNNSGKDNTSANSADSE.....KNLSRAKEMLTNSVKELTS 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Swlssprot\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	658.5	36.4	210	1	OSCL_BORBU 007337 borrelia bu
2	556	30.7	212	1	OSCL_BORBU 008137 borrelia bu
3	392.5	21.7	214	1	VM24_BORHE P22778 borrelia he
4	373	20.6	215	1	VM03_BORHE Q02448 borrelia he
5	181	10.0	1957	1	YD86_SCHPO Q00799 plasmodium
6	180.5	10.0	1251	1	RBP2_PLAVB P25386 saccharomyc
7	178.5	9.9	1790	1	USOL_YEAST P19401 streptococ
8	173	9.6	564	1	M12_STRPY P05655 acanthamoeb
9	171.5	9.5	1509	1	MTSN_ACACA P32399 bacillus su
10	169.5	9.4	775	1	YHGE_BACSU P12845 caenorhabdi
11	169.5	9.4	1947	1	MYSC_CAEEL P13535 homo sapien
12	168	9.3	1937	1	MYH8_HUMAN P02566 caenorhabdi
13	167	9.2	1966	1	MYSB_CAEEL Q16787 homo sapien
14	166.5	9.2	1713	1	LMAS3_HUMAN Q00796 plasmodium
15	165.5	9.1	2869	1	RBP1_PLAVB Q04139 rickettsia
16	165	9.1	998	1	SCA4_RICAK P17739 borrelia bu
17	164.5	9.1	296	1	OSBI_BORBU Q28641 oryctolagus
18	164.5	9.1	1938	1	MYH4_RABIT P30622 homo sapien
19	163	9.0	1427	1	REST_HUMAN P02565 gallus galli
20	163	9.0	1940	1	MYH3_CHICK P08089 streptococ
21	162.5	9.0	483	1	M6_STRPY Q42184 gallus galli
22	162.5	9.0	1433	1	REST_CHICK P43454 mesocricetu
23	162.5	9.0	3210	1	CENF_HUMAN Q94137 rickettsia
24	161.5	8.9	1526	1	MYH2_SCHPO P02562 oryctolagus
25	161.5	8.9	1941	1	MYH2_HUMAN P43237 saccharomyc
26	161	8.9	981	1	SCA4_RICEE P13540 mesocricetu
27	161	8.9	1084	1	MYSS_RABIT P12883 homo sapien
28	160.5	8.9	1938	1	MYSS_RABIT P02562 oryctolagus
29	160	8.8	679	1	MYSS_YEAST P43237 saccharomyc
30	159	8.8	1934	1	MYH7_MESAU P13540 mesocricetu
31	159	8.8	1935	1	MYH7_HUMAN P12883 homo sapien
32	157.5	8.7	1939	1	MYH1_HUMAN P12883 homo sapien
33	157	8.7	1186	1	SMC_BACSU P18134 bacillus su

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	210 AA.
OSCL_BORBU	007337			
AC	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB19.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-ATCC 35210 / B31;			
RC	MEDLINE-93239332; PubMed-8478108;			
RX	Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E., Schwab E., Wanner G.;			
RA	"Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of Borrelia burgdorferi.";			
RT	Infect. Immun. 61:2182-2191(1993).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-ATCC 35210 / B31;			
RC	MEDLINE-96025162; PubMed-7494039;			
RX	Fukunaga M., Hamase A.;			
RA	"Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu lato isolates from Japan.";			
RT	J. Clin. Microbiol. 33:2415-2420(1995).			
RL	[5]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-ATCC 35210 / B31;			
RC	MEDLINE-96025162; PubMed-7494039;			
RX	Freder C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., Iacharya R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., Dougherty B., Tomb J.-F., Fleischman R.D., Richardson D., Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			

34	157	8.7	1935	1	MYH7_PIG	P79293 sus scrofa
35	155.5	8.6	2349	1	TPR_HUMAN	P12270 homo sapien
36	155	8.6	1102	1	MYSC_CHICK	P29616 gallus galli
37	154.5	8.5	465	1	MYH8_RABIT	P04460 oryctolagus
38	154.5	8.5	2017	1	MYSN_DROME	O99323 drosophila
39	154.5	8.5	2116	1	MYH2_DICDI	P08799 dictyosteli
40	153.5	8.5	1167	1	CAGA_HELPI	O94111 helicobacte
41	153	8.5	705	1	YNP9_CAEEL	P34562 caenorhabdi
42	153	8.5	1013	1	SCA4_RICRH	O96181 rickettsia
43	153	8.5	1935	1	MYH7_RAT	P02564 rattus norv
44	151.5	8.4	539	1	MYH3_HYDAT	P39922 hydra atten
45	151.5	8.4	1939	1	MYH4_HUMAN	O94623 homo sapien

RA Uterback T., Matthey L., McDonald L., Artlisch P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "genomic sequence of a Lyme disease spirochaete, *Borrelia*  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X69596; CAA49306.1; -  
 EMBL: U01894; AA16058.1; -  
 EMBL: D49497; BAA08457.1; -  
 DR EMBL: AE000792; AAC66329.1; -  
 DR TIGR: BBH19; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 KW Complete proteome.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 210 OUTER SURFACE PROTEIN C.  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 SQ SEQUENCE 210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match 36.4%; Score 658.5; DB 1; Length 210;  
 Best Local Similarity 76.7%; Pred. No. 5.7e-23;  
 Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

OY 1 MACNNSGKDGNTSANSDESVKGPNTLEISKTTSDSNAVLAVKEVALLSIDEL-ARK 59  
 DB 17 ISCNNGSGKDGNTSANSDESVKGPNTLEISKTTSDSNAVLAVKEVALLSIDEL-ARK 76  
 OY 60 ICKKINDGSLNEANRNESLAGAVTSTLTQKLSKNGSGLEKTAARKCEERS 119  
 DB 77 ICKKIQNNGLDTENNHNSGLLAGAVTSTLTQKLDGK-NEGLEKIDAAKCESETF 135  
 DB 120 TKLKDHQAOLGIGYTDENAKKAILKANAGKDGVELEKLSGLESLSKAKKEMLAN 179  
 DB 136 NKLEKHTDLGKRGVTDADAKKAILKTNGT-KTGAELGLKLFESVEVLAKAKKEMLAN 194  
 OY 180 VKELTSPV 188  
 DB 195 VKELTSPV 203  
 RESULT 2  
 OSC2\_BORBU STANDARD; PRT; 212 AA.  
 ID OSC2\_BORBU  
 AC 008137;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE OUTER SURFACE PROTEIN C PRECURSOR (PC).  
 GN OSCP.  
 OS *Borrelia burgdorferi* (Lyme disease spirochaete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
 OX NCBI\_TaxID=139;  
 RN (1)  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=PKO;  
 RX MEDLINE=92219995; PubMed=1560779;  
 RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

RA Soutschek E.;  
 RT "Molecular analysis and expression of a *Borrelia burgdorferi* gene  
 RT encoding a 22 kDa protein (PC) in *Escherichia coli*.";  
 RL Mol. Microbiol. 6:503-509(1992).  
 RN [2]  
 RP SEQUENCE OF 1-205 FROM N.A.  
 RC STRAIN=DK26;  
 RX MEDLINE=94075528; PubMed=8253951;  
 RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
 RT "Polymorphism in ospC gene of *Borrelia burgdorferi* and  
 RT immunoreactivity of OSCP protein: implications for taxonomy and for  
 RT use of OSCP protein as a diagnostic antigen.";  
 RL J. Clin. Microbiol. 31:2570-2576(1993).  
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X62162; CAA44093.1; -  
 EMBL: X73624; CAA52003.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 KW Complete proteome.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 212 OUTER SURFACE PROTEIN C.  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 SQ SEQUENCE 212 AA; 22499 MW; C206C31FBF2E7D4 CRC64;

Query Match 30.7%; Score 556; DB 1; Length 212;  
 Best Local Similarity 67.2%; Pred. No. 1.7e-18;  
 Matches 123; Conservative 19; Mismatches 39; Indels 2; Gaps 2;

OY 191 NNSGKDGNT-SANSDESVKGPNTLEISKTTSDSNAVLAVKEVALLSIDEL-ARK 248  
 DB 20 NNSGKDGSDASTNPADSESKGPNLEISKTTSDSNAVLAVKEVALLSIDEL-ARK 79  
 OY 249 KKKKNQVSDNDNAGSILAGAVTSTLTQKLSKNGSGLEKTAARKCEERS 308  
 DB 80 OKIDNNGSLALNNONGSLAGAVTSTLTQKLSKNGSGLEKTAARKCEERS 139  
 OY 309 LGEHTDLGKRGVTDENAKKAILKTNGT-KTGAELGLKLFESVEVLAKAKKEMLAN 368  
 DB 140 LKSGHADLGKQATDHAHAAILKTHTATDKAKKEDLFESVEVLAKAQAVALNSVKE 199  
 OY 369 LTS 371  
 DB 200 LTS 202

RESULT 3  
 VM24\_BORHE STANDARD; PRT; 214 AA.  
 ID VM24\_BORHE  
 AC P32778;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-OCT-1993 (Rel. 27, Last annotation update)  
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.  
 GN VMP24.  
 OS *Borrelia hermslii*.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
 OX NCBI\_TaxID=140;  
 RN (1)  
 RP SEQUENCE FROM N.A.



RC STRAIN-SSP. HSI SEROTYPE 24;  
 RA MEDLINE-93133110; PubMed-1484486;  
 RT Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids  
 are highly polymorphic."  
 RT Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L04786; AAA22964.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KM Outer membrane; Lipoprotein; Signal; Plasmid.  
 FT SIGNAL 1 18 PROBABLE.  
 FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT SEQUENCE 214 AA: 22541 MW: F1583F510246FC7 CRC64;

Query Match 21.7%; Score 392.5; DB 1; Length 214;  
 Best Local Similarity 47.6%; Pred. No. 2.4e-11;  
 Matches 90; Conservative 32; Mismatches 54; Indels 13; Gaps 3;

OY 191 NNSGKDGNTSANSADSVKGP---NLTEISKITTESNAVLAVKEVELLTSIDELAKA 246  
 DB 20 NNGGPE-----LKSDEVAKSDGTVDLAKVSKKIKKSAFASVKEVELTVKSVDELAKA 74  
 OY 247 IGGKIKNDVS-LDNEADHNGSLISGAYLISNLTTRKISAID---SGELKAETEKAKKC 302  
 DB 75 IGGKIKNDGDLDEAGONSLLAGVSVSAVIRKVALETTSIGISNELTKITEVSKSA 134  
 OY 303 SEFTAKLKEHNDLKEGVTDDNAKRAIKLTNNDKTKGAELEKLEESVKNLSKAKEM 362  
 DB 135 EAFNLKLGKGTHTLGGKKAASDDTKKAIKKNDSKTKGASLEALMTAVDALAKAAGEV 194  
 OY 363 TNSVKELETS 371  
 DB 195 EAARKELTA 203

RESULT 4  
 YD86\_SCHPO STANDARD: PRT; 215 AA.  
 AC 002448;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.  
 GN Borrelia hermsli.  
 OS Borrelia hermsli.  
 OC Plasmid.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=140;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SSP. HSI SEROTYPE 3;  
 RX MEDLINE-93133110; PubMed-1484486;  
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

RT are highly polymorphic."  
 RL Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP24.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L04789; AAA22967.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KM Outer membrane; Lipoprotein; Signal; Plasmid.  
 FT SIGNAL 1 18 PROBABLE.  
 FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT SEQUENCE 215 AA: 23139 MW: 684C74D35F87C771 CRC64;

Query Match 20.6%; Score 373; DB 1; Length 215;  
 Best Local Similarity 46.3%; Pred. No. 1.7e-10;  
 Matches 88; Conservative 32; Mismatches 56; Indels 14; Gaps 4;

OY 191 NNSGKDGNTSANSADSVKGP---NLTEISKITTESNAVLAVKEVELLTSIDELAKA 246  
 DB 20 NNGGPE-----LKSDEVAKSDGTVDLAKVSKKIKKSAFASVKEVELTVKSVDELAKA 74  
 OY 247 IGGKIKNDVS-LDNEADHNGSLISGAYLISNLTTRKISAID---SGELKAETEKAKKC 301  
 DB 75 IGGKIKNDNPNFEDENHNGSLIAGVSVILTVKAKLTSLDIOITIGISDELTKTEGVAKKE 134  
 OY 302 SEFTAKLKEHNDLKEGVTDDNAKRAIKLTNNDKTKGAELEKLEESVKNLSKAKEM 361  
 DB 135 SEAFVTVQVSKKHTDLKAEVTAHAASAILVDTGDKGAELIKLTAFIDELAKAANDA 194  
 OY 362 LTNSVKELETS 371  
 DB 195 VETVIRKELTA 204

RESULT 5  
 YD86\_SCHPO STANDARD: PRT; 1957 AA.  
 AC 010411;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.  
 GN SPAC1F3.06C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 270690; CAA94624.1; -  
KW Hypothetical protein.  
FT NON\_TER 1  
SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 10.0%; Score 181; DB 1; Length 1957;  
Best Local Similarity 23.8%; Pred. No. 0.47;  
Matches 110; Conservative 74; Mismatches 149; Indels 130; Gaps 23;

QY 8 KQGNISANSADK---SVKGNLFEISKITDSNAVLAV--KEVEALLSSIDE----- 55  
DB 360 KDSRTNSQLEEMVWELKSNRT--HSQTLDAESKLSFQENKSKSGSIDQNNLSK 418  
DB 56 -----LKAIGK-----KIKND-----GSLDNEA 74  
DB 419 DKWQVSSQLEEARSSLAHATGKLAENSERDPQKKIKDEKIRBQDRACINSSNEL 478  
QY 75 NNESLLAGAYITLTLQKLSKNGSEGLKIAAKKCEEFSTKL-----KD 124  
DB 479 KEK-----SALIDKQDELN--NLKQIKQKQVSESTOSLSQIDRIINEKK 525  
QY 125 NH-----AOLG-IQG-----VTDENAKKAILKANAAKDKGV---FELEKLSGSLSLSKA 171  
DB 526 KHEVYESQNELKGELOTEISNHSQSLTSLAEKEAAVATNNELSKSKSLQTLCA 585  
QY 172 AKEMLANSKELTSPVYHGNNSGKQGNISANSADSVKGNLFEISKITDSN--AVLAV 230  
DB 566 FEKELAKSVWOL-----KENQNFSSLDTSYFK--KLINSHQELNNHQTITKOL 632  
QY 231 KEVEITLTSIDELAKAIGKIKNDVSLDNE-----ADHNGSLISGAYILSLITRK 281  
DB 633 KOTSSKLOOL-QLERANPEQKSTSLSDENNDLRTKLLEESNKSILKQEDVDL--EKN 690  
QY 282 ISAIKDSGELKAEIKAKKCS---EEFTAKLGEHTDLGKE-----GVTDNNAKAIL 331  
DB 691 IOTLKEDLRKSEALRFSLKLEAKNREVYDNLKQKHETLEAQGRNDLHSLDAKNTNAIL 750  
QY 332 KTNNDKTKGADLEKLFESYKML---SKAAKEMLTNSKELTS 371  
DB 751 --SSELTSSSEDKRRLTANVTLLTQDSKAMQSFSTLVNSYOS 791

LT 6  
RBP2\_PLAVB  
ID RBP2\_PLAVB STANDARD; PRT; 1251 AA.  
AC Q00799;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).  
GN RBP2.  
OS Plasmodium vivax (strain Belen).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=31273;  
RN 11  
RP SEQUENCE FROM N.A.  
RA MEDLINE=92313338; PubMed=1617731;  
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
RT "A reticulocyte-binding protein complex of Plasmodium vivax  
merozoites";  
RT Cell 69:1213-1226(1992).  
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
CC HUMAN RETICULOCYTE CELLS.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).  
CC CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M88098; AAA29744.1; -  
KW Malaria; Receptor; Membrane.  
FT NON\_TER 1  
SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 10.0%; Score 180.5; DB 1; Length 1251;  
Best Local Similarity 21.8%; Pred. No. 0.3;  
Matches 104; Conservative 63; Mismatches 160; Indels 151; Gaps 21;

QY 5 NSGKGNIS--ANSADSVKGNLFEISKITDSNAVLAV-----REVEALLSIDEL 56  
DB 670 NEG-HGNVQQLVENIKELVDENNLSDILKQATGRKEEIOKITHSTLKKAKATLIGHVDT 728  
QY 57 AKAIKKI-----KNDGSLDNE--ANNR-----SLAGA 84  
DB 729 AKYVGKITPPELATELLGDAKLTAQELKEPKSKNNVYLETEMNKRNTNEDVHKNIODA 788  
QY 85 YITSTLI-----TQKLSKNGSEGL-----KEKIAAKKCEEFSTKL- 122  
DB 789 YKVALEIILAHSEIDITRKQDSKLIEMGNQIYLKVVLYINQKNKISSIKSEAVSAKIG 848  
QY 123--KDNHAGLQIGVTDENAKKAIL-----KANAAKDKGVELEKLSG 163  
DB 849 NVSKRHSLESLKTCSDKSYDNIHLEKQTELONLNSFTQEKTYNTNSDK---LEKIKT 904  
QY 164 SLESSTAKAKEMLANSVKELTSPVYHGNNSGKQGNISANSADSVKGNLFEISKITDS 223  
DB 905 DPESLKAKLKTLEGE-----VNALKASSDNEHVQSSSEPV--NALSEIKET-- 952  
QY 224 NAVVLAVKEVELTSLSIDELAK-----AIGKKIKNDVSLDNEADHNGSLI 268  
DB 953 -----DISLNTALDELAKKGTCEVSRVYKLIKQVTEIIDDITLITIEKNK-- 1002  
QY 269 SGAYL-----ISMLT--KKISAIKDSGELKAEIKAKKCEEFSTAKLGEHTDL 316  
DB 1003 --AVLAYIKRNEPTVOVLTLNHFNTKQVSNHEPTNFKNSKSEBELTKAVYDSKTI 1060  
QY 317 GKEGVTDNNAKAILIKTN-----NDKTRKGADELKLFESYKNSLAKAKEM--LNSVK 367  
DB 1061 SK-----LKGVIIEVENTEMNTIESSAKETLALINELNKKRTSLNEIYQTSNEVK 1111

RESULT 7  
USOL\_YEAST  
ID USOL\_YEAST STANDARD; PRT; 1790 AA.  
AC P25386;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL1.  
GN USOL1 OR INT1 OR YDI058W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN 11  
RP SEQUENCE FROM N.A.  
RA STRAIN=X2180-1A;  
RA MEDLINE=91185402; PubMed=2010462;  
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
RA Yamasaki M.;  
RT "A cytoskeleton-related gene, usol1, is required for intracellular  
RT protein transport in Saccharomyces cerevisiae";  
RT J. Cell Biol. 113:245-260(1991).  
RN 12  
SEQUENCE OF 782-1790 FROM N.A.

RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,  
RA Kendrick R.E.;  
RL Submitted (FEB-1993) to the EMBL/Genbank/DBD databases.  
RN [3]  
RP SEQUENCE OF 1-8 FROM N.A.  
RA Bai Y., Symington L.S.;  
RL Submitted (MAY-1996) to the EMBL/Genbank/DBD databases.  
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
CC COMPLEX.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
CC ER AND THE GOLGI COMPLEX.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEATITIVE, COMPOSED  
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X54378; CAA38253.1; -  
DR EMBL: L03188; AAB00143.1; -  
DR EMBL: U53668; AAB6659.1; -  
DR PIR: A38455; A38455.  
DR HSSP: P80240; 1DIP.  
DR SGD: S0002216; USO1.  
DR InterPro: IPR002017; Spectrin.  
KM Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
FT DOMAIN 1 724  
FT DOMAIN 725 1790  
FT DOMAIN 465 487  
FT DOMAIN 991 1790  
FT DOMAIN 1172 1786  
FT DOMAIN 847 847  
FT CONFLICT 924 924  
FT CONFLICT 1253 1253  
FT CONFLICT 1319 1319  
FT CONFLICT 1461 1461  
FT CONFLICT 1581 1581  
FT CONFLICT 1600 1600  
FT CONFLICT 1661 1661  
FT CONFLICT 1772 1772  
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;  
Query Match 9.9%; Score 178.5; DB 1; Length 1790;  
Best Local Similarity 23.7%; Pred. No. 0.54;  
Matches 89; Conservative 68; Mismatches 156; Indels 63; Gaps 13;  
OY 44 KEVALLSSIDELAKAIGKKI-KNDGSLNEANRNESLLAGVATITSLTLTKQSKLN--- 99  
DB 1022 KNEIOLAKTISDLEFOTKEETISKSDSKDEYEQISLLEKLEKLETTATTANDENNNKISL 1081  
OY 100 -GSEGLIKERIAAKKCSEEFSTKLKDNHQAOLGIQVTDENAKKAILKANAGKQVVEL 158  
DB 1082 KTRRELEAEALAYKNKLELEKLETSKALKEVENEHLKREKIQLEKEARET-KQOL 1140  
OY 159 EKLSGSLESLK-----AA-----KEMLANSVKELTSPVYHGNSKDKGTSANSD 208  
DB 1141 NSLRNLSELEKEHEDLAOLKQKYEIOJANKKEOYNEEISQLN-----DETSTQOENESI 1196  
OY 209 KGPN-----LLEISKRTTESNAVYLAKEV-----ETLTSIDELAKAIGKK 250  
DB 1197 KKKNDLEBEVAVAMKSTSEQSNLKSSEIDALNLOIKELKKKNETEASLLESIKSEVE 1256  
OY 251 IKNDVSLDNEADHNGSLGAVLISNLITRKISAIKDSG-----ELKAETEKAKKCSEEF 306  
DB 1257 TVKIKLOEDCNFKEKEVE-----LFDKILKASEDKNSKYLELQKESSEKIKIEELDARK 1309

OY 307 AKLK-----GERTDGLKRGVTDNNAKKAITKNNNDTKGAD-LEKL-----PESVK 352  
DB 1310 TELKIQLEKITTNLNSKAKESSELSRLKTSSEKRNKAEQLKNEIKNQAFKER 1369  
OY 353 NLSKAKEMLTNSVKE 368  
DB 1370 KLINEGSSITTOEYSE 1385  
RESULT 8  
M12 STRPY STANDARD: PRT: 564 AA.  
AC P19401;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).  
GN EMM12.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CS24 / Serotype M12;  
RX MEDLINE=88058777; PubMed=2445730;  
RA Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;  
RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream  
sequences.";  
RL J. Bacteriol. 169:5633-5640(1987).  
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES  
CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF  
CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO  
CC PHAGOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.  
CC -1- SIMILARITY: TO OTHER M PROTEINS.  
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS  
CC IN THE REGION OF THE MEMBRANE ANCHOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M18269; AAB88573.1; -  
DR PIR: A60115; A60115.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR003345; M\_repeat.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR Pfam: PF02370; M; 9.  
DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
KM Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
KM Transmembrane; Coiled coil; Signal.  
FT SIGNAL 1 41  
FT CHAIN 42 564  
FT DOMAIN 42 550  
FT TRANSMEM 551 564  
FT DOMAIN 44 505  
FT DOMAIN 505 541  
FT DOMAIN 542 547  
FT NON\_TER 564 564  
SQ SEQUENCE 564 AA; 62904 MW; 5F1549DCA7B46 CRC64;  
Query Match 9.6%; Score 173; DB 1; Length 564;  
Best Local Similarity 23.1%; Pred. No. 0.26;  
Matches 100; Conservative 56; Mismatches 149; Indels 128; Gaps 19;

```

OY 10 GNTSANSADSVKPNLITETSKITDSNAVLAVKEVEALLSSIDELAKAI--GKKIKND 67
DB 115 GKIGINDADLAK--ITTELEKSVSEKNOVLSQIKK-----ELEAEADIQGRVYH-- 163
OY 68 GSIDNANRRESILAGAVITITITK-----LSKLINGS--EGLKEKIAAKKCEEFSTKUK 123
DB 164 -AADLLRHQOE-----IAKEENVISKLNGELOPLKQKVDETRNLOQEKQVYL 210
OY 124 DNHAQIGIOGVDEKNAKAILKANANGKDKGVLELEKSGSLESLSKAAKEMLANSVKL 183
DB 211 SLEQOL--AVYKENAKKPELA-ALGHOLADKEYNAKIAELSES-----KLADAKKDF 259
OY 184 TSPVY--HGNV-----SGKDG-----NTSANSADSVKGP----- 211
DB 260 ELAALGHQAHNREYQAKLEKDKQIQKQLEQKQIUDASKRGAROLEAVRAKKAETAE 319
OY 212 -NLTELSKITESNAVLAVK-----EVEITLITIDE--LAKAIGKKI 251
DB 320 NMLKAEIARVTEOKQIILDSRKGTARDLEAVRKSQKQVEALAKOLEQNKISEASRKGL 379
OY 252 KNDVSLDNEADHNGSLISGAYLISNLITKKSIAIKDSGSELKAEIKAKCSEFTAKLKG 311
DB 380 RRDLOTSRA-----KQVEKDLANLITELDKVKEKOISDASRQK 420
OY 312 EHTDLGKEGVTDNNAKAILKLTNDKTKGADLEKLFESVKNL-----SK 356
DB 421 LRRLDASREAKQYKEA-LEBANSKILAEKLNKDLSESKLTTEREKAELQAKLEAEK 479
OY 357 AAKEMLTNSVKEL 369
DB 480 ALKEQALAKQAEEL 492

RESULT 9
ID MYSN ACACA STANDARD: PRT: 1509 AA.
AC P05659.
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS Acanthamoeba castellanii (Ameoba).
OC Eukaryota, Acanthamoebidae, Acanthamoeba.
OX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=87308395; PubMed=3040773;
RT Hammer J.A., III, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT hinge in the rodlike tail."
RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CHARACTERS OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00624; CAA68663.1;
DR PIR: A27224; A27224.
DR HSSP: P08799; IAMD.
DR InterPro: IPRO00048; IQ.
DR InterPro: IPRO01609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW Methylation; Alkylation; Phosphorylation; Multigene family.
FT DOMAIN 1 789
FT DOMAIN 848 1509
FT DOMAIN 848 1509
FT DOMAIN 1227 1509
FT DOMAIN 1253 1509
FT DOMAIN 1253 1482
FT DOMAIN 1483 1509
FT NP_BIND 182 189
FT DOMAIN 660 682
FT DOMAIN 766 780
FT MOD_RES 133 133
FT MOD_RES 700 700
FT MOD_RES 1489 1489
FT MOD_RES 1494 1494
FT MOD_RES 1499 1499
SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173017E CRC64;

Query Match 9.5%; Score 171.5; DB 1; Length 1509;
Best Local Similarity 24.3%; Pred. No. 0.91;
Matches 97; Conservative 70; Mismatches 143; Indels 89; Gaps 18;

OY 19 ESKVGNLITETSKITDSNAVLAVKEVA---LLSSIDELAKAIGKKIKNDGSLDNEA 74
DB 953 EAERGEELKASLEEEERNRKALQOEAKTKVESERNELQOEYDEAAHDSLKKEEDLSREL 1012
OY 75 NRNSILAGAVITITLTOKLSK-NSSEGLK---EKIAAKKCEEFSTKLDNNAQLG 130
DB 1013 RETDALADENISSETLRSKIKTERGADVRLNLDVTAFTKLOETKKSLEEEIAQTR 1072
OY 131 IOGVTDENAKAILKANANGKDKGVLELE-----KLSGLESLSKAAKEMLAN 179
DB 1073 AQLEEEKSGKEA---ASSKAKQIG-QQLEPARSEVDSLKLSNAEKSL-KTAKDQNRDL 1127
OY 180 VKELTSPVYHGNNSGKQKNTSANSADSVKGP--PNLITETSKIT--ESNAVLAVKEY 233
DB 1128 DEOL-----EDERTVRANVDKQKALPAKLELTDQVATLDGQKNA---AAQA 1173
OY 234 ETLITSIDEL-----AKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKIS 283
DB 1174 KILKTOYDETKRRLREAEASAAIRLEKERKN--ALDEVAQ-----LTAIDL 1216
OY 284 AAKDSG-----ELKAEIERAKK---CSEFTAKLNGEHTDAGKEGVTDNNAK 328
DB 1217 AERDSGAQQRKRLNTRISLEQSELENAPKIGASSEVVKRLBELRLBELILLTAQEA 1276
OY 329 AILKTNDRKKGADLEKLFESVKNLSKAAKEMLTNSVK 367
DB 1277 AAEK-NIDKAN--LELEELRQEADDAARDNDKLVKNRK 1312

RESULT 10
ID YHGE_BACSU STANDARD: PRT: 775 AA.
AC P32399;

```

DT 01-OCT-1993 (Rel. 27, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HYPOTHECAL 84.1 KDA PROTEIN IN HEMV-GLTT INTERGENIC REGION (ORFB).  
 GN YHGE.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Bacillus.  
 NX NCBI\_TaxId=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-145 FROM N.A.  
 RX MEDLINE=93094140; PubMed=1459957;  
 RA Hanson M., Hederstedt L.;  
 RT "Cloning and characterization of the Bacillus subtilis hemv gene  
 RT cluster, which encodes proteoheme ix biosynthetic enzymes."  
 RI J. Bacteriol. 174:8081-8093(1992).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: TO L. LACTIS PHAGE INFECTION PROTEIN (PIP).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: Y14083; CAAT4522.1; -;  
 DR EMBL: M97208; AAA22521.1; -;  
 DR EMBL: Z99109; CAB12856.1; -;  
 DR PIR: F47045; F47045.  
 DR Subtilast: BG10433; yHGE.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 16 36 POTENTIAL.  
 FT TRANSMEM 585 605 POTENTIAL.  
 FT TRANSMEM 625 645 POTENTIAL.  
 FT TRANSMEM 655 675 POTENTIAL.  
 FT TRANSMEM 742 762 POTENTIAL.  
 SO SEQUENCE 775 AA; 84093 MW; 49A5A6A6A70B66 CRC64;

Query Match 9.4%; Score 169.5; DB 1; Length 775;  
 Best Local Similarity 21.5%; Pred. No. 0.53;  
 Matches 91; Conservative 74; Mismatches 141; Indels 117; Gaps 19;

5 NSGKD--GNTSANSADSVGPNLITEISKITDSNAVLAVKEVALSLSDLAKEAI-- 60  
 135 MAGSNVYGATIGEKAIKDKL-----ASYSKEVTE-----QYTKVLFDEKDKAKGLSD 182  
 DB 135 MAGSNVYGATIGEKAIKDKL-----ASYSKEVTE-----QYTKVLFDEKDKAKGLSD 182  
 QY 61 ---GKRIKNGSLDNEANRRESLLAGAVTITLITOKLSKNGSEGEKIAAKKCSSE 117  
 DB 183 ASSGAKKIDGKTDAK-----NSAQLEKNIATKAKESTAT 217  
 QY 118 FSTRK--LKNHQAQL--GIO-----GVTDENAKKAILKAN--AAGKDKGVEELEKSGS 164  
 DB 218 ISDKTQALADGAQAVTSLSLGGKPDSSNOIYDKSQAAGSGELTSKKNELLAG 277  
 QY 165 LESLSRAAKMLANSVKELTSPVYHGNNSGKDGNTSANSADSVGPNLITEISKITDSN 224  
 DB 278 LQWVQKTPN-LTNGLDQLNSKV-----QEGSEKAAKKEIKINADLTLETAVNMLE 329  
 QY 225 AVYLAWEVETLTSLDELAKGKIKINDVSLDN--EADHNGSLSGA----- 271  
 DB 330 KSETAKMEFKQLDFENSLKNRDAFKNTNSDPLTADEKSQLINSVEKKLPDVPADP 389  
 QY 272 --VLISNLIT---KKISAIKDSGE--LKAETIEKAKKCSSEFTAKLKEHNDLAGEVYTD 324  
 DB 390 FDOITSLQPLPADQPLDIATIKSSLEEDVKAQVAVQ--KAMPEATSKLY----- 434

QY 325 NAKAIIKTNNDKTKGADL-----EKLEPSYKNL-----SKAKEMLTNSVK 367  
 DB 435 NGAITIDALDRLTEGADKTYNGSKLTDGOTKLTAGIGEVNKPAPAKAGSEDLVTGSS 494  
 QY 368 ELT 370  
 DB 495 QVS 497

RESULT 11  
 MYSC\_CAEEL STANDARD; PRT; 1947 AA.  
 AC P12845;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MYOSIN HEAVY CHAIN C (MHC C).  
 GN MYO-2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;  
 OC Rhabdilitidae; Peloderinae; Caenorhabdilitis.  
 NX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=89178677; PubMed=2926820;  
 RA Dibb N.J., Maruyama I.N., Krause M., Karn J.;  
 RT "Sequence analysis of the complete Caenorhabdilitis elegans myosin  
 RT heavy chain gene family."  
 RL J. Mol. Biol. 205:603-613(1989).  
 RN [2]  
 RP SEQUENCE OF 1-21; 112-371 AND 1501-1772 FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=85201409; PubMed=3888374;  
 RA Karn J., Dibb N.J., Miller D.M.;  
 RT "Cloning nematode myosin genes."  
 RL Cell Muscle Motil. 6:185-237(1985).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MYOSIN HEAVY CHAIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE PHARYNGEAL MUSCLE.  
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.  
 CC ELEGANS.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X08066; CAAT30855.1; -;  
 DR EMBL: M37233; AAA28121.1; -;  
 DR EMBL: M37235; AAA28122.1; -;  
 DR EMBL: M37236; AAA28123.1; -;  
 DR PIR: S05697; S05697.  
 DR HSP: P08799; LMND.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.

DR pfam: PF00063; myosin\_head; 1.  
 DR pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00242; MYSC; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Multigene family.  
 FT DOMAIN 1 855 MYOSIN HEAD-LIKE.  
 FT DOMAIN 856 1947 ROD-LIKE TAIL (S2 AND LMM DOMAINS).  
 FT NP\_BIND 174 181 COILED COIL (POTENTIAL).  
 FT DOMAIN 668 690 ATP (BY SIMILARITY).  
 FT DOMAIN 773 787 ACTIN-BINDING.  
 FT MOD\_RES 125 125 METHYLATION (TR1-) (POTENTIAL).  
 FT MOD\_RES 708 708 ALKYLATION (SH-1).  
 FT MOD\_RES 718 718 ALKYLATION (SH-2).  
 FT CONFLICT 132 132 E -> D (IN REF. 2).  
 FT CONFLICT 137 137 M -> I (IN REF. 2).  
 SEQUENCE 1947 AA; 223009 MW; 6D5AE99BD013627 CRC64;

Query Match 9.4%; Score 169.5; DB 1; Length 1947;  
 Best Local Similarity 22.9%; Pred. No. 1.5; Indels 135; Gaps 18;  
 Matches 93; Conservative 61; Mismatches 118;

QY 23 GPNLLEISRTDSNAVLLAVKEVALLSSIDELAKAIGKIKKNDG-----SLD 71  
 DB 1202 GQDLTVLKKGSDA-----IQLSPQIEQLQKQRIKKGKMGHOREDESCALD 1252  
 QY 72 NNA--NRNSLLAGATYST-----LITQKSLKNGS-----EGIK 105  
 DB 1253 QAKRLADDERLAKGEVYTSERLRKADQSHOLDPVSCK--GRINSENDSRLARVEEIE 1311  
 QY 106 EKIAAKKCESEFSTRKKNHQAOLGIGVTDENAKKAILKANAAAGKDGVELEKLSGL 165  
 DB 1312 AKIQANRLKLOFENEL--DHAKRAEESRRKRLSNLSKILA-----RELEQLKESI 1363  
 QY 166 E-----SLSKAKEMLANSVKELTSPVHGNSGKDNTSANSADSVKPN-- 212  
 DB 1364 EDEVAGKNEASRQLSKASVELDQWRTKEET-----EGLIGADPEDEKVRKQNK 1412  
 QY 213 LTEISKKITESAVYLAKEVEITLTSIDELAKAIGKIKKNDVSDNADNHSLSICAY 272  
 DB 1413 TSEIDALDACNAKIVLEMAERSRLT-----AEADANRL-----FAEHHAQVSS-- 1457  
 QY 273 LISNLTFRKISAI-----KDSGELKAETEKAKKCESEFTA---KKEGERTDGGKGV 321  
 DB 1458 -----LEKKQAFKFDVIDEMKKKVDVLYELDGAQRDARQLSGEAKHLRGQHTL----- 1507  
 QY 322 TDNAKKAAILKTNDKTGADLEKLFESVKNLSKAKEMLTNSVKE 368  
 DB 1508 -----ADQVEGLRREKNKSLSDETRD-LTESLSE 1534

RESULT 12  
 MYH8\_HUMAN STANDARD; PRT; 1937 AA.  
 AC P13535; Q14910;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MYOSIN HEAVY CHAIN, SKELETAL MUSCLE, PERINATAL (MYHC-PERINATAL).  
 GN MYH8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90323631; PubMed=2373371;  
 RA Karsch-Mizrachi I., Peghali R., Shows T.B. Jr., Leinwand L.A.;  
 "Generation of a full-length human perinatal myosin heavy-chain-

RT encoding cDNA.";  
 RL Gene 89:289-294(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=95324556; PubMed=7601129;  
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Stedman H.H., Rubinstein N.A.;  
 RT "Characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RL Eur. J. Biochem. 230:1001-1006(1995).  
 RN [3]  
 RP SEQUENCE OF 502-1937 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90235862; PubMed=1691980;  
 RA Bober E., Buchberger-Seidel A., Braun T., Singh S., Goedde H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 myosin heavy chains.";  
 RL Eur. J. Biochem. 189:55-65(1990).  
 RN [4]  
 RP SEQUENCE OF 860-1937 FROM N.A.  
 RX MEDLINE=89234168; PubMed=2715179;  
 RA Peghali R., Leinwand L.A.;  
 RT "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RL J. Cell Biol. 108:1791-1797(1989).  
 RN [5]  
 RP SEQUENCE OF 1-46 FROM N.A.  
 RA Esser K., Tichar A., Myszkowski M.;  
 RT "Isolation and characterization of the human perinatal MYC promoter.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M36769; AAC17185.1; -;  
 DR EMBL: Z38133; CA86293.1; -;  
 DR EMBL: X51592; CA83594.1; -;  
 DR EMBL: M35250; AAA36346.1; -;  
 DR EMBL: AF067143; AAC21557.1; -;  
 DR PIR: A30220; A30220.  
 DR HSSP: P13538; 2MYS.  
 DR MIM: 160741; -;  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR002298; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.

DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PSS0096; IQ; 1.  
 KM Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM ATP-binding; Methylation; Alkylation; Multigene family;  
 KM Calmodulin-binding.  
 FT DOMAIN 1 780 MYOSIN HEAD-LIKE.  
 FT DOMAIN 842 1937 COILED COIL (POTENTIAL).  
 FT NP\_BIND 181 188 ATP.  
 FT DOMAIN 658 680 ACTIN-BINDING.  
 FT DOMAIN 760 774 ACTIN-BINDING.  
 FT MOD\_RES 132 132 METHYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 698 698 ALKYLATION (SH-2) (POTENTIAL).  
 FT MOD\_RES 708 708 ALKYLATION (SH-2) (POTENTIAL).  
 FT CONFLICT 15 15 A -> R (IN REF. 2).  
 FT CONFLICT 970 970 E -> Q (IN REF. 1 AND 4).  
 FT CONFLICT 1072 1072 M -> N (IN REF. 3).  
 FT CONFLICT 1247 1247 N -> H (IN REF. 1 AND 4).  
 FT CONFLICT 1251 1252 MC -> DGG (IN REF. 3).  
 FT CONFLICT 1261 1261 E -> G (IN REF. 1 AND 4).  
 FT CONFLICT 1287 1287 K -> Q (IN REF. 1 AND 4).  
 FT CONFLICT 1377 1378 KY -> Q (IN REF. 3).  
 FT CONFLICT 1504 1505 EN -> AH (IN REF. 1 AND 4).  
 FT CONFLICT 1847 1847 E -> D (IN REF. 1 AND 4).  
 FT CONFLICT 1914 1914 D -> H (IN REF. 2).  
 SQ SEQUENCE 1937 AA: 222762 MW: 438251517929E8 CRC64;

Query Match 9.3%; Score 168; DB 1; Length 1937;  
 Best Local Similarity 23.2%; Pred. No. 1.7;  
 Matches 88; Conservative 70; Mismatches 166; Indels 56; Gaps 14;

QY 15 NSADESVKPNLTETSKKTTDSNAVLAVK-----EVEALLSIDELAKAIKKINDGS 69  
 DB 914 NKIQLEAKIKETVERAEFEIEINAEIETAKRKLEDCSELEKIDIDLETLAKVEKEKA 973  
 QY 70 LNEARNRNSLAGATITLTITQKSLKNGSGLEKIAAKKCEPESTKIKDNHQL 129  
 DB 974 TEKKVKN-----LTEEMA-----GLOETIA--KLSEK--KKAQETHQQT 1009  
 QY 130 GIGCVTDENAKKAILKANAAGKDGVEELEKLSLESLSKAKEMLANSEKELTPVH 189  
 DB 1010 LDDLOAEKKVNLTKA-----KTKLEQYDDLEGSLEQEKKLMD--LERAKKRLGLEGLK 1064  
 QY 190 GNNSGKDGNTSANSADSV--KGNLTETSKKTTESNAVLAV--KEYETLLTIDEL--- 243  
 DB 1065 AOSTMDMENDKQDLEKLEKFEISNLISKIEDQAVEIQLOKKIKELQARIEELGEE 1124  
 QY 244 -----AKAIKKIKNDVSLD--NEADHNGSLISGAVLISLTKKISAIKDSGLAEITE 296  
 DB 1125 IEAERASRAKAEKORSDELRELEIEISERLEEAGATSAOVELKKRRA--EPEKLRDLE 1182  
 QY 297 KAKKCEPEFAKJKEGHTD--IGKEGYTDNAKKAAILKTNDKTKGADLEKLEPESVKNLS 355  
 DB 1183 EATLQHEAMVAALAKKHADSMALGEOIDMLQRYKOKLEKSELMETDLSNNAEATS 1242  
 QY 356 KAKEM-----LTNSVKEL 369  
 DB 1243 KAKGNLEKKCRSLDQVSEL 1262

RESULT 13  
 MYSB\_CAEEL  
 ID MYSB\_CAEEL STANDARD: PRT; 1966 AA.  
 AC P02566;  
 DT 21-JUL-1986 (rel. 01, Created)  
 DT 21-JUL-1986 (rel. 01, Last sequence update)  
 DT 20-AUG-2001 (rel. 40, Last annotation update)  
 DE MYOSIN HEAVY CHAIN B (MHC B).  
 GN UNC-54 OR MYO-4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83273600; PubMed=6576334;  
 RA Karn J., Brenner S., Barnett L.;  
 RT "Protein structural domains in the Caenorhabditis elegans unc-54  
 myosin heavy chain gene are not separated by introns."  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).  
 RN [2]  
 RP SEQUENCE OF 850-1966 FROM N.A.  
 RX MEDLINE=8227395; PubMed=7202124;  
 RA McLachlan A.D., Karn J.;  
 RT "Periodic charge distributions in the myosin rod amino acid sequence  
 match cross-bridge spacings in muscle."  
 RL Nature 299:226-231(1982).  
 RN [3]  
 RP SEQUENCE OF 1876-1966 FROM N.A.  
 RX MEDLINE=8323892; PubMed=6571695;  
 RA Willis N., Gesteland R.F., Karn J., Barnett L., Bolten S.,  
 RA Waterston R.H.;  
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber  
 nonsense mutations via altered transfer RNA."  
 RL Cell 33:575-583(1983).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE  
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.  
 ELEGANS.  
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY  
 WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: J01050; AAA28124.1; -  
 DR EMBL: V01494; CAA24738.1; -  
 DR PIR: A02992; MKWK  
 DR HSSP: P08799; 1MND  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR Pfam: PF01576; Myosin\_tail.1.  
 DR PRINTS: PR00193; MYOSINHEAVY  
 DR PRODOM: PD000355; myosin\_head.1.  
 DR SMART: SM00242; MYSC; 1.  
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Multigene family.  
 KM ATP-binding; Methylation; Alkylation; Multigene family.  
 FT DOMAIN 1 850 COILED COIL (POTENTIAL).  
 FT DOMAIN 851 1966 ALPHA-HELICAL TAILPIECE (S2).  
 FT DOMAIN 851 1166 HINGE.  
 FT DOMAIN 1165 1176 LIGHT MEROMYOSIN (LMW).  
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMW).  
 FT NP\_BIND 177 184 ATP (BY SIMILARITY).  
 FT DOMAIN 665 687 ACTIN-BINDING.  
 FT DOMAIN 769 783 ACTIN-BINDING.







FT	CHAMIN	21	1713	LAMININ ALPHA-3 CHAIN.
FT	DOMAIN	46	201	DOMAIN IIT A.
FT	DOMAIN	67	185	2.5 X LAMININ EGF-LIKE REPEATS.
FT	DOMAIN	114	113	LAMININ EGF-LIKE 1.
FT	DOMAIN	167	166	LAMININ EGF-LIKE 2.
FT	DOMAIN	202	793	LAMININ EGF-LIKE 3 (INCOMPLETE).
FT	DOMAIN	794	1713	DOMAIN II AND I (HEPAT REPEATS).
FT	DOMAIN	794	970	5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT	DOMAIN	971	1139	LAMININ G-LIKE 1.
FT	DOMAIN	1140	1353	LAMININ G-LIKE 2.
FT	DOMAIN	1354	1529	LAMININ G-LIKE 3.
FT	DOMAIN	1354	1713	LAMININ G-LIKE 4.
FT	DOMAIN	1330	1713	LAMININ G-LIKE 5.
FT	DOMAIN	231	327	COILED COIL (POTENTIAL).
FT	DOMAIN	396	548	COILED COIL (POTENTIAL).
FT	DOMAIN	594	621	COILED COIL (POTENTIAL).
FT	DOMAIN	702	765	COILED COIL (POTENTIAL).
FT	DOMAIN	1686	1713	COILED COIL (POTENTIAL).
FT	DISULFID	67	76	BY SIMILARITY.
FT	DISULFID	69	83	BY SIMILARITY.
FT	DISULFID	86	95	BY SIMILARITY.
FT	DISULFID	98	111	BY SIMILARITY.
FT	DISULFID	114	126	BY SIMILARITY.
FT	DISULFID	116	135	BY SIMILARITY.
FT	DISULFID	137	146	BY SIMILARITY.
FT	DISULFID	149	164	BY SIMILARITY.
FT	DISULFID	202	202	INTERCHAIN (PROBABLE).
FT	DISULFID	205	205	INTERCHAIN (PROBABLE).
FT	SITE	658	660	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	542	542	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	645	645	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	745	745	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	882	882	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	964	964	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1108	1108	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1131	1131	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1325	1325	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1477	1477	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1667	1667	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	VANSPPLIC	1	45	MGMLTFEALALDQCLGYSQQQKRVFPLDPGQSLQASIVE
FT				FRPS -> KVSYSYGYLTYQAKSFALPDQWVLEKRPVQL
FT				TCQHSLSIYETETNPRLRHGRHVVHVEGNFRSHSRAPV
FT				SREELMTVSLRADVRIGQLVFTETQRLLSVSGLEASDT
FT				GSGRALAVETACAPRAYAGDSC (IN ISOBORM B).
FT	CONFLICT	5	5	W -> R (IN REF. 2).
FT	CONFLICT	123	125	ATG -> GMC (IN REF. 2).
FT	CONFLICT	481	481	M -> K (IN REF. 2).
FT	CONFLICT	754	754	R -> L (IN REF. 2).
FT	CONFLICT	969	969	E -> Q (IN REF. 2).
FT	CONFLICT	1052	1052	D -> A (IN REF. 2).
FT	CONFLICT	1184	1184	G -> A (IN REF. 2).
FT	SEQUENCE	1713 AA:	189304 MW:	45EA9BEB017B60D3 CRC64;

Query Match	9.2%	Score 166.5	DB 1	Length 1713
Best Local Similarity	23.0%	Pred. No. 1.7		
Matches	90	Conservative	71	Mismatches 138; Indels 93; Gaps 16
QY	9	DGNTSANSADSESVKGPNTLEISKTTDSNAVL	---	LAVEYBALLSIDELAKAIGKRI 64
DB	396	ENNGLANSTIRDS	---	LNEYEAKISDLRLAQEAAQAQKANGINGLQEMERALGATQGV 450
QY	65	KNDGSLDWEANNESLGAATITSLITQKLSLNGSELKRIKIAA	-	KKCEEPSTRIK 123
DB	451	KTINSLOSFTYUETTAODSSLQOTINIAQLMK	---	SOKEYEKLAASLVEAQOELSDDYR 507
QY	124	DHNAOLGIQVTDENAKKAILKANAAGKDGVEELEKLSGLESLSKAKEMIAN	-	SVYE 182
DB	508	E	---	LSRSAGKSTLVEBAEKHARSQDELAKOJEBELIKRNASGDE 547
QY	183	LN	---	SPVYHGNSSGKDGNTSANSADSESVKGPNTLEISKTTESNAVYLAWE 232
DB	548	LVRCAVDAATAYENTLNLAKAEDANRAARASSES	---	ALQTVIVE 590

ID	RPPL_PLAVB	STANDARD	PRT	2869 AA.
AC	Q00798			
DT	01-APR-1993	(Rel. 25, Created)		
DT	01-APR-1993	(Rel. 25, Last sequence update)		
DT	01-OCT-1996	(Rel. 34, Last annotation update)		
DE	RETICULOCYTE BINDING PROTEIN 1 PRECURSOR.			
GN	RBPL1.			
OS	Plasmodium vivax (strain Belen).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
ON	NCBI_Taxid:31273;			
OX	11			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE:92315338; PubMed:1617731;			
RA	Galinaki M.R., Medina C.C., Ingravallo P., Barnwell J.W.;			
RT	"A reticulocyte-binding protein complex of Plasmodium vivax merozoites".			
RT	Cell 65:1213-1226(1992).			
CC	-1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO			
CC	HUMAN RETICULOCYTE CELLS.			
CC	-1- SUBUNIT: HOMODIMER (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: M88097; AAA29743.1; .			
DR	HSSP: P36956; IAM9.			
KW	Malaria; Receptor; Signal; Transmembrane.			
FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	2869	RETICULOCYTE BINDING PROTEIN 1.
FT	DOMAIN	18	2807	EXTRACELLULAR.
FT	TRANSMEM	2808	2826	POTENTIAL.
FT	DOMAIN	2827	2869	CYTOPLASMIC.
FT	SITE	1030	1032	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	2599	2601	CELL ATTACHMENT SITE (POTENTIAL).
SO	SEQUENCE	2869 AA;	330213 MW;	B9DBE442205EBCF CRC64;

```

Query Match Score 9.1%; DB 1; Length 2869;
Best Local Similarity 22.4%; Pred. No. 3.4;
Matches 101; Conservative 63; Mismatches 124; Indels 163; Gaps 20;

OY 25 NITETSKIT-DSNAVLAKEVEALLSSIDELAKAIGKIKKIDGSLDNEANRNESLAG 83
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 1141 HNTVAAGHTSKRNKILSVKEDKRLNVEDQ-NEDRYKKVNP--ENE----- 1185
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

OY 84 AVTISTLTOKRSTKNGS-EGLEKRIAAAKCSEEFSTRKLNHQAOLGIGQVTDENAKKA 142
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 1186 -----KQLEAIRGSKSRKLEYI-----NKH-----VSEKTOLESTAN 1217
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

OY 143 IUKANAAGKDR--GYEELKTLGSL-----ESTLSKAKEMLANVSKELTSPVYHGNNSGD 196
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

```
Db 1218 TLKSNAGKENEHDEELNKTGOMRDIYEKUKIAEELKGTYNEL-----KD 1266
Oy 197 GNTSANSADSVKGPNTLEISKIT-ESNAVYLAWEVETLLTSIDELAK-----245
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1267 ANEKANKVEPEPERNIGHVLERITVEKDKAGKVEEMNSLKTIEKLIQETSDSQNEL 1326
Oy 246 ---AIGKKIKN-----DVSLDNEAD-----HNGSLISGA 271
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1327 VTTSTKHLNNAKGYEDVTKRNEEDSIQIREKAKSLETLDENKKLVQOVNMNLQSAIQN 1386
Oy 272 YLISN-----LITKKISAI-----KDSGELK---AE 294
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1387 AGISKELNELKGVIELLISTNNSILIEYKKNSSSVRFSSOLANGEFTKABEGEKNASAR 1446
Oy 295 IEKAKKSEEFPAKLKGEHTDLGKEGYTDDNAKKAALK-----TNNDKTK-----339
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1447 LAEAEKLEQIYKDL--DYSIDIDKVKIEGIRKREILKMKESALTFWEESEKFKOMCSSH 1504
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
    340 --GADELEKLFESVKMLSKAKEMLTNSVKE 368
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1505 MENAKEGKKRIEYLNKNGDGKANITDSOME 1535
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: March 18, 2002, 10:11:52  
Job time: 970 sec



QY 181 KELTSPPV 188  
 DB 187 KELTSPPV 194

## RESULT 2

044977 PRELIMINARY; PRT; 211 AA.  
 AC 044977;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C.  
 GN OSCP.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PARE:  
 RX MEDLINE=95395018; PubMed=7655660;  
 RA Jauris-Helpe S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,  
 RA Soutschek E., Will G., Wilske B.;  
 RT "Molecular analysis of genes encoding outer surface protein C (OspC)  
 RT of Borrelia burgdorferi sensu lato: relationship to ospa genotype and  
 RT evidence of lateral gene exchange of ospC.";  
 RL J. Clin. Microbiol. 33:1860-1866(1995).  
 DR EMBL; X81522; CA57242.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 SQ SEQUENCE 211 AA; 22214 MW; BFE8C7820CDE10D3 CRC64;

Query Match 49.0%; Score 886; DB 2; Length 211;  
 Best Local Similarity 97.3%; Pred. No. 1.7e-32;  
 Matches 183; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVALLSIDELAKAI 60  
 DB 17 ISCNNGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVALLSIDELAKAI 76  
 QY 61 GKIKNDGSLDNEANRNESILAGAVTISTLTOKSLKNGSEGLKETAANKCEEST 120  
 DB 77 GKIKNDGSLDNEANRNESILAGAVTISTLTOKSLKNGSEGLKETAANKCEEST 136  
 QY 121 KIKDNHAIQIGVTDENAKKAILKANAAGKGVLELEKLSGLSLSKAKEMLANSV 180  
 DB 137 KIKDNHAIQIGVTDENAKKAILKANAAGKGVLELEKLSGLSLSKAKEMLANSV 196  
 QY 181 KELTSPPV 188  
 DB 197 KELTSPPV 204

RESULT 3  
 031120 PRELIMINARY; PRT; 191 AA.  
 ID 031120;  
 AC 031120;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSCP.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OC10;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
 RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF029869; AAB86552.1;

DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 191 AA; 20502 MW; E4FB56BCB61740F9 CRC64;

Query Match 48.9%; Score 885; DB 2; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-32;  
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 NNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVETLTSIDELAKAI 250  
 DB 10 NNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVETLTSIDELAKAI 69  
 QY 251 IKNDVSLDNEADHNSLSLISGAVTISLTOKSLKNGSEGLKETAANKCEESTAKL 310  
 DB 70 IKNDVSLDNEADHNSLSLISGAVTISLTOKSLKNGSEGLKETAANKCEESTAKL 129  
 QY 311 GEHTDLGREGVTDNNAKKAIIKTNDKTKGADLEKLFESYKNTSKAKEMLTNSVKELT 370  
 DB 130 GEHTDLGREGVTDNNAKKAIIKTNDKTKGADLEKLFESYKNTSKAKEMLTNSVKELT 189  
 QY 371 S 371  
 DB 190 S 190

RESULT 4  
 P94242 PRELIMINARY; PRT; 193 AA.  
 ID P94242;  
 AC P94242;  
 DT 01-MAY-1997 (TREMblrel. 03, Created)  
 DT 01-MAY-1997 (TREMblrel. 03, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSCP.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ZS7;  
 RX MEDLINE=96296448; PubMed=8709845;  
 RA Lavey I., Gibbs C.P., Schuster R., Dorner F.;  
 RT "Evidence for lateral transfer and recombination in OSCP variation in  
 RT Lyme disease Borrelia.";  
 RL Mol. Microbiol. 18:257-269(1995).  
 DR EMBL; I42868; AAB37011.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 193 AA; 20189 MW; 64E30C240E9F54CA CRC64;

Query Match 48.9%; Score 884; DB 2; Length 193;  
 Best Local Similarity 98.4%; Pred. No. 1.9e-32;  
 Matches 183; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVALLSIDELAKAI 62  
 DB 1 CNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVALLSIDELAKAI 60  
 QY 63 KIKNDGSLDNEANRNESILAGAVTISTLTOKSLKNGSEGLKETAANKCEESTKTL 122  
 DB 61 KIKNDGSLDNEANRNESILAGAVTISTLTOKSLKNGSEGLKETAANKCEESTKTL 120  
 QY 123 KDNHAIQIGVTDENAKKAILKANAAGKGVLELEKLSGLSLSKAKEMLANSV 182  
 DB 121 KDNHAIQIGVTDENAKKAILKANAAGKGVLELEKLSGLSLSKAKEMLANSV 180

OY 183 LTSPV 188  
DB 181 LTSPV 186

## RESULT 5

OSPC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID-139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-297;  
RA MEDLINE-96296448; PubMed-8709845;  
RX Lacey I., Gibbs C.P., Schuster R., Dorner F.;  
RT "Evidence for lateral transfer and recombination in ospC variation in  
Lyme disease Borrelia".  
RL Mol. Microbiol. 18:257-269(1995).  
DR EMBL; L42893; AAB37001.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT SEQUENCE 192 AA; 20472 MW; 46AC8F93E4DFED6C CRC64;

Query Match 48.6%; Score 879; DB 2; Length 192;  
Best Local Similarity 99.4%; Pred. No. 3.1e-32;  
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 191 NNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAIGKK 250  
DB 2 NNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAIGKK 61  
OY 251 INKDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELKAEIKAKKCSSEFTAKLK 310  
DB 62 INKDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELKAEIKAKKCSSEFTAKLK 121  
OY 311 GEHTDLGKEGVTDDNKKALIKTNNDKTKGADLEKLFESVKNLSRAKEMLTNSVKELT 370  
DB 122 GEHTDLGKEGVTDDNKKALIKTNNDKTKGADLEKLFESVKNLSRAKEMLTNSVKELT 181  
OY 371 S 371  
DB 182 S 182

## RESULT 6

OSPC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID-139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HB19;  
RA MEDLINE-95154673; PubMed-7651744;  
Stevenson B., Barthold S.W.;

RT "Expression and sequence of outer surface protein C among North  
American isolates of Borrelia burgdorferi."  
RL FEMS Microbiol. Lett. 124:367-372(1994).  
DR EMBL; U04281; AAC43297.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT SEQUENCE 210 AA; 22526 MW; 2D672991D584E4EE CRC64;

Query Match 48.6%; Score 879; DB 2; Length 210;  
Best Local Similarity 99.4%; Pred. No. 3.4e-32;  
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 191 NNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAIGKK 250  
DB 20 NNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAIGKK 79  
OY 251 INKDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELKAEIKAKKCSSEFTAKLK 310  
DB 80 INKDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELKAEIKAKKCSSEFTAKLK 139  
OY 311 GEHTDLGKEGVTDDNKKALIKTNNDKTKGADLEKLFESVKNLSRAKEMLTNSVKELT 370  
DB 140 GEHTDLGKEGVTDDNKKALIKTNNDKTKGADLEKLFESVKNLSRAKEMLTNSVKELT 199  
OY 371 S 371  
DB 200 S 200

RESULT 7  
Q08138 PRELIMINARY; PRT; 203 AA.

AC Q08138;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).  
GN OSPC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID-139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DK7;  
RX MEDLINE-94075528; PubMed-8253951;  
RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
RT "Polymorphism in ospC gene of Borrelia burgdorferi and  
immunoreactivity of OspC protein: implications for taxonomy and for  
use of OspC protein as a diagnostic antigen".  
RL J. Clin. Microbiol. 31:2570-2576(1993).  
CC -1 SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
ANCHOR.  
DR EMBL; X73625; CAA52004.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
KW Signal; Lipoprotein; Outer membrane.  
FT SIGNAL 1 18  
FT CHAIN 19 203  
FT NON\_TER 203  
FT SEQUENCE 203 AA; 21364 MW; AC9B29FE0E69B48F CRC64;

Query Match 47.7%; Score 862; DB 2; Length 203;  
Best Local Similarity 95.7%; Pred. No. 1.9e-31;  
Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAI 60  
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAVKEVETLLTSIDELAKAI 76

```

RA MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RA "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RA U.S.A.'?";
RA Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP
RP SEQUENCE FROM N.A.
RC STRAIN=TE7S;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91795; AAB81892.1; -.
DR Interpro: IPR001800: Lipoprotein_6.
DR Pfam: PF01441: Lipoprotein_6; 1.
DR ProDom: PD001149: Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 174
FT NON_TER 174
SQ SEQUENCE 174 AA; 18291 MW; A14F5B8BA151A278 CRC64;

Query Match 43.7%; Score 790; DB 2; Length 174;
Best Local Similarity 97.0%; Pred. No. 2.3e-28;
Matches 162; Conservative 3; Mismatches 2; Indels 0; Gaps

OY 1 MACNNSGKDGNTSNASDESVKGPNTLETISKRTSDSNAYLVKREYVALLSIDELAKAI 60
DB :|||||
DB 8 ISCNNSGKDGNTSNASDESVKGPNTLETISKRTSDSNAYLVKREYVALLSIDELAKAI 67
OY 61 GKRTIKNGSDLDNEANNRSESLLAGAVYITSLITQKLSKNGSEGLKEKIAAKKCEEST 120
DB :|||||
DB 68 GKRTIKNGSDLDNEANNRSESLLAGAVYITSLITQKLSKNGSEGLKEKIAAKKCEEST 127
OY 121 KLKNHNAOLGIGVTDENAKKAILTKNNAKGKKGVELEKLSGSIES 167
DB 128 KLKNHNAOLGIGVTDENAKKAILTKNNAKGKKGVELEKLSGSIES 174

RESULT 10
O9AGBI ID O9AGBI PRELIMINARY; PRT; 165 AA.
AC O9AGBI:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OSCP-MODIFIED.
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=HB19;
RA Kumaran D., Esvaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
RA Lawson C.L., Swaminathan S.;
RA "Crystal structure of outer surface protein C (oscp) from the Lyme
RA disease spirochete, Borrelia burgdorferi.";
RL EMBL J. 0:0-0(2001).
DR EMBL; AF337548; AAK21289.1; -.
FT VARIANT 61 M -> I.
FT VARIANT 174
SQ SEQUENCE 165 AA; 17833 MW; 56B01536D22F61BF CRC64;

Query Match 43.4%; Score 786; DB 2; Length 165;
Best Local Similarity 98.2%; Pred. No. 3.2e-28;
Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps

OY 208 VKPNTLETISKRTESNAVYLVKREYVLLSIDELAKIGKRTKNDYSIDNEADHNGSI 267
DB :|||||
DB 1 MKPNTLETISKRTESNAVYLVKREYVLLSIDELAKIGKRTKNDYSIDNEADHNGSI 60
OY 268 ISGAVYISNLITRKISAIKDSGELKAETKAKKCEESTAKLKGHEHDLGKEGVTDNNK 327
DB :|||||
DB 61 MSGAVYISNLITRKISAIKDSGELKAETKAKKCEESTAKLKGHEHDLGKEGVTDNNK 120

```

OY 328 KALIKTNDKGADELEKLFESVKNLSKAKEMLTNSVKELTS 371  
 DB 121 KALIKTNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELTS 164

RESULT 11  
 034121 PRELIMINARY: PRT: 168 AA.

AC 034121: 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TETS;  
 RX MEDLINE=97478003: PubMed=9336916;  
 RA Ras N.M., Postic D., Foretz M., Baranton G.;  
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
 U.S.A.'?";  
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TETS;  
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U91794; AAB81891.1; -;  
 DR InterPro: IPR001800: Lipoprotein\_6.  
 DR Pfam: PF01441: Lipoprotein\_6; 1.  
 DR ProDom: PD001149: Lipoprotein\_6; 1.  
 FT NON\_TER 1 168  
 FT SEQUENCE 168 AA; 17704 MW; CDA45DD8A49C4409 CRC64;

Query Match 43.2%; Score 782; DB 2: Length 168;  
 Best Local Similarity 98.2%; Pred. No. 5e-28;  
 Matches 164; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 19 ESKVGNLLEISKRTDSNVAVLLAVEVALLSIDELAKIKKTKNDSDNEANRNE 78  
 DB 1 ESKVGNLLEISKRTDSNVAVLLAVEVALLSIDELAKIKKTKNDSDNEANRNE 60  
 OY 79 SLTAGVTTSTLTITOKLSKNGSEGLKEKTAIAKKCESEFSTKLKDNHQAOLGIGVTDEN 138  
 DB 61 SLTAGVTTSTLTITOKLSKNGSEGLKEKTAIAKKCESEFSTKLKDNHQAOLGIGVTDEN 120  
 OY 139 AKKAILKANAAKDKGVEELEKLSGLESLSKAKEMLANSVKELTS 185  
 DB 121 AKKAILKANAAKDKGVEELEKLSGLESLSKAKEMLANSVKELTS 167

RESULT 12  
 031115 PRELIMINARY: PRT: 193 AA.  
 AC 031115: 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OC3;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF029862; AAB86545.1; -;

DR InterPro: IPR001800: Lipoprotein\_6.  
 DR Pfam: PF01441: Lipoprotein\_6; 1.  
 DR ProDom: PD001149: Lipoprotein\_6; 1.  
 FT NON\_TER 1 193  
 FT SEQUENCE 193 AA; 20596 MW; 018A4CB310475A58 CRC64;

Query Match 39.8%; Score 720.5; DB 2: Length 193;  
 Best Local Similarity 79.8%; Pred. No. 2.9e-25;  
 Matches 150; Conservative 16; Mismatches 21; Indels 1; Gaps 1;

OY 1 MACNNSGKDNSTANSDESCKGNLLEISKRTDSNVAVLLAVEVALLSIDELAKAI 60  
 DB 7 ISCNNSGKDNASNSADESVKGNLLEISKRTDSNVAVLLAVEVTTLASIDELAKAI 66  
 OY 61 GKIKKNDGSLDNEANRNEISLAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCESEFST 120  
 DB 67 GKIKKNDVSLDNEADNNGSLIAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCESEFTK 126  
 OY 121 KLDNHAQOLGIGVTDENAKKAILKANAAKDKGVEELEKLSGLESLSKAKEMLANSV 180  
 DB 127 KLEKHTDGLCKKDATDVHAKAELIKTNGT-KDKGALEKLFESVENLAKAEMLSNV 185  
 OY 181 KELTSPVY 188  
 DB 186 KELTSPVY 193

RESULT 13  
 031122 PRELIMINARY: PRT: 194 AA.  
 AC 031122: 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OC12;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF029871; AAB86554.1; -;  
 DR InterPro: IPR001800: Lipoprotein\_6.  
 DR Pfam: PF01441: Lipoprotein\_6; 1.  
 DR ProDom: PD001149: Lipoprotein\_6; 1.  
 FT NON\_TER 1 194  
 FT SEQUENCE 194 AA; 20640 MW; 00A5E6E2D2CE0F7F CRC64;

Query Match 36.7%; Score 664; DB 2: Length 194;  
 Best Local Similarity 75.1%; Pred. No. 9.1e-23;  
 Matches 142; Conservative 19; Mismatches 26; Indels 2; Gaps 2;

OY 1 MACNNSGKDNSTANSDESCKGNLLEISKRTDSNVAVLLAVEVALLSIDELAKAI 59  
 DB 7 ISCNNSGKDNSTANSDESCKGNLLEISKRTDSNVAVLLAVEVTTLASIDELAKAI 66  
 OY 60 GKIKKNDGSLDNEANRNEISLAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCESEFST 119  
 DB 67 IGKTIQNGGLAVDAGNGTLAGAVTTSTLTITOKLGLKNSSEKLEIAKKCSDDFT 126  
 OY 120 TKLDNHAQOLGIGVTDENAKKAILKANAAKDKGVEELEKLSGLESLSKAKEMLANS 179  
 DB 127 KLEGEHAQOLGIENVTDENAKKAILITDAA-KDKGALEKLFKAVENLAKAEMLSNV 185  
 OY 180 VKELTSPVY 188  
 DB 186 VKELTSPVY 188

DB 186 VKETSPV 194

RESULT 14

ID 034120 PRELIMINARY; PRT; 184 AA.

AC 034120;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI\_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-TETS;

RX MEDLINE=97478003; PubMed=9336916;

RA Ras N.M., Postic D., Foretz M., Baranton G.;

RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the U.S.A.'?"

RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-TETS;

RA Marti-Ras N., Postic D., Foretz M., Baranton G.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U91793; AAB81890.1; "

DR InterPro: IPR001800; Lipoprotein\_6.

DR Pfam: PF01441; Lipoprotein\_6; 1.

DR Prodom: PD001149; Lipoprotein\_6; 1.

FT NON-TER 1

FT NON-TER 184

SEQUENCE 184 AA; 19391 MM; 9B1C12B79273FB0E CRC64;

Query Match 36.6%; Score 662; DB 2; Length 184;

Best Local Similarity 76.3%; Pred. No. 1e-22;

Matches 142; Conservative 15; Mismatches 27; Indels 2; Gaps 2;

OY 3 C NNSGKDGNTSANSADSVKGNLTETISKITDSNAVLLAVEVALLSSIDELAKAIG 62

DB 1 C NNSGKDGNTSANSADSVKGNLTETISKITDSNAVLLAVEVALLSSIDELAKAIG 60

OY 63 KIKNDGSLDNEANRNESILAGAYTISLTITOKLSKNGSEGLKEKIAAKKSEEFSTRK 122

DB 61 KVEANG-LGNEDRRTSLLAGHEISILITOKLTALKDSGGLKAEIARKKCEAFTRKL 119

OY 123 KDNHAQIGIGVTDENAKKAILKANAAAGKDGVELEKLSGSLSELSKRAKEMLANSVK 182

DB 120 KDNHAQIGIGVTDENAKKAILKANAAAGKDGVELEKLSGSLSELSKRAKEMLANSVK 178

OY 183 LTSPV 188

DB 179 LTNPV 184

RESULT 15

P94234

ID P94234 PRELIMINARY; PRT; 193 AA.

AC P94234;

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS OSpC.

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI\_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-28354;

RX MEDLINE=96296448; PubMed=8709845;

RA Livey I., Gibbs C.P., Schuster R., Dörner F.;

RT "Evidence for lateral transfer and recombination in OSpC variation in Lyme disease Borrelia."

RT Mol. Microbiol. 18:257-269(1995).

RL EMBL: L42895; AAB37003.1; "

DR InterPro: IPR001800; Lipoprotein\_6.

DR Pfam: PF01441; Lipoprotein\_6; 1.

DR Prodom: PD001149; Lipoprotein\_6; 1.

FT NON-TER 1

FT NON-TER 193

SEQUENCE 193 AA; 20411 MM; 05B68720F061E2A0 CRC64;

Query Match 36.6%; Score 662; DB 2; Length 193;

Best Local Similarity 75.9%; Pred. No. 1.1e-22;

Matches 142; Conservative 17; Mismatches 26; Indels 2; Gaps 2;

OY 3 C NNSGKDGNTSANSADSVKGNLTETISKITDSNAVLLAVEVALLSSIDELAKAIG 61

DB 1 C NNSGKDGNTSANSADSVKGNLTETISKITDSNAVLLAVEVALLSSIDELAKAIG 60

OY 62 KIKNDGSLDNEANRNESILAGAYTISLTITOKLSKNGSEGLKEKIAAKKSEEFSTRK 121

DB 61 KKIQNGGLAVEAGHNGTLTLAGAYTISLTITOKLIDGLKNSEKLEKIAAKKSEDFTRK 120.

OY 122 KDNHAQIGIGVTDENAKKAILKANAAAGKDGVELEKLSGSLSELSKRAKEMLANSVK 181

DB 121 LEGBHAQIGIGVTDENAKKAILITDAA-KDKGAELERLFRVAENLAKAKEMLANSVK 179

OY 182 LTSPV 188

DB 180 ELTSPV 186

Search completed: March 18, 2002, 10:10:54

Job time: 977 sec



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:32 ; Search time 118.14 Seconds  
(without alignments)  
230.734 Million cell updates/sec

Title: us-09-596-746a-38

Perfect score: 1797  
Sequence: 1 MACNNNGKDGKNGASNSADES.....ENLSKAADTLKNAVKELTS 368

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

```
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1797	100.0	368	22	AA62717
2	1789	99.6	391	22	AA62736
3	1533	85.3	367	22	AA62719
4	1530	85.1	391	22	AA62735
5	1485.5	82.7	368	22	AA62718
6	1482.5	82.5	392	22	AA62734
7	1479	82.3	369	22	AA62714
8	1476	82.1	393	22	AA62732
9	1225.5	68.2	377	22	AA62713
10	1224	68.1	373	22	AA62711
11	1221	67.9	397	22	AA62729

12	1220	67.9	369	22	AA62716	Borrelia sp chimera
13	1219.5	67.9	400	22	AA62739	Borrelia sp chimera
14	1217	67.7	393	22	AA62731	Borrelia sp chimera
15	1178	65.6	378	22	AA62712	Borrelia sp chimera
16	1176.5	65.5	374	22	AA62710	Borrelia sp chimera
17	1173.5	65.3	398	22	AA62728	Borrelia sp chimera
18	1172.5	65.2	370	22	AA62715	Borrelia sp chimera
19	1172	65.2	401	22	AA62738	Borrelia sp chimera
20	1169.5	65.1	394	22	AA62730	Borrelia sp chimera
21	1167.5	65.0	560	22	AA62724	Borrelia sp chimera
22	1157.5	64.4	384	22	AA62726	Borrelia sp chimera
23	1154.5	64.2	408	22	AA62737	Borrelia sp chimera
24	1142	63.6	378	22	AA62725	Borrelia sp chimera
25	1134	63.1	401	22	AA62733	Borrelia sp chimera
26	1099.5	61.2	410	22	AA62740	Borrelia sp chimera
27	1098.5	61.1	386	22	AA62727	Borrelia sp chimera
28	901	50.1	191	22	AA62705	B burgdorferi ospc
29	899	50.0	184	22	AA62702	B burgdorferi ospc
30	889	50.0	191	22	AA60884	Borrelia 28691 O
31	889	49.5	191	22	AA62706	B burgdorferi ospc
32	673	37.5	466	16	AA675740	B31 outer surface
33	663	36.9	193	22	AA62709	B burgdorferi ospc
34	657.5	36.6	212	16	AA611934	Outer surface prot
35	651.5	36.3	194	15	AA60885	Borrelia 2591 Ospc
36	640	35.6	587	16	AA675746	B31 Osp-A/antigen
37	638.5	35.5	466	16	AA675739	B31 outer surface
38	623.5	34.7	192	22	AA62703	B burgdorferi ospc
39	623.5	34.7	210	16	AA611935	Outer surface prot
40	622.5	34.6	192	22	AA62721	B burgdorferi ospc
41	621.5	34.6	192	15	AA60886	Borrelia ip2 Ospc
42	620.5	34.5	209	22	AA62720	B burgdorferi ospc
43	617.5	34.4	210	16	AA675727	B. burgdorferi str
44	608.5	33.9	178	15	AA62771	Borrelia 2591 anti
45	605.5	33.7	194	15	AA60894	Borrelia H9 antige

## ALIGNMENTS

```
RESULT 1
ID AAB62717 standard; Protein: 368 AA.
XX
AC AAB62717;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SPQ ID NO: 38.
XX
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN W0200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Sehnost G, Dykhuizen D, Luft RJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
XX
DR N-PSDB; AAF29021.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
```

XX Claim 43; Page 97; 160pp; English.

PS  
CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX

XX Sequence 368 AA;

Query Match 100.0%; Score 1797; DB 22; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.5e-107;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSGKDNASANSADSVKGPNTLTKITSSNAVYLAKEVEETLLASIDELATKA 60  
DB 1 macnnsqkdnasansadesvkgpnlteiskkltesnavylavkevetllasidelatka 60  
QY 61 IGGKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTEFTMK 120  
DB 61 igkigngngleangskntsllsgayaissdilaeklnvlkneelkekidtakgscstefmk 120  
QY 121 LKSEHAVLGLDNLTDNAORAILKKHANKDKGAELKLFKAVENTLSKAAODTLKNAYKE 180  
DB 121 lksehavlgldnltdnagrailkkhankdkgaalekfkavenlskaagdtlknayke 180  
QY 181 LTPSYHGNNSRKDNASNSADESVKGPNTLTKITSSNAVYLAKEVEETLLASIDE 240  
DB 181 ltpsyhgnnsrkdnasnsadesvkgpnlteiskkltesnavylavkevetllaside 240  
QY 241 LATKAIIGKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOST 300  
DB 241 latkaiigkigngngleangskntsllsgayaissdilaeklnvlkneelkekidtakgscst 300  
QY 301 EFTNKLSEHAVLGLDNLTDNAORAILKKHANKDKGAELKLFKAVENTLSKAAODTLK 360  
DB 301 eftnklsehavlgldnltdnagrailkkhankdkgaalekfkavenlskaagdtlkn 360  
QY 361 NAVKELTS 368  
DB 361 navkeltss 368

RESULT 2

AAB62736  
ID AAB62736 standard; Protein: 391 AA.

XX AAB62736;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 76.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PA (UYNV ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29040.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 144-145; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX

XX Sequence 391 AA;

Query Match 99.6%; Score 1789; DB 22; Length 391;  
Best Local Similarity 99.7%; Pred. No. 5.1e-107;  
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDNASANSADSVKGPNTLTKITSSNAVYLAKEVEETLLASIDELATKAI 61  
DB 25 scnnsgkdnasansadesvkgpnlteiskkltesnavylavkevetllasidelatka 84  
QY 62 GKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTEFTMK 121  
DB 85 gkligngngleangskntsllsgayaissdilaeklnvlkneelkekidtakgscstefmk 144  
QY 122 KSEHAVLGLDNLTDNAORAILKKHANKDKGAELKLFKAVENTLSKAAODTLKNAYKE 181  
DB 145 ksehavlgldnltdnagrailkkhankdkgaalekfkavenlskaagdtlknayke 204  
QY 241 TSPYHGNNSRKDNASNSADESVKGPNTLTKITSSNAVYLAKEVEETLLASIDE 241  
DB 241 tspyhgnnsrkdnasnsadesvkgpnlteiskkltesnavylavkevetllaside 264  
QY 264 ATKAIIGKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTE 301  
DB 265 atkaiigkigngngleangskntsllsgayaissdilaeklnvlkneelkekidtakgscst 324  
QY 302 FTNKLSEHAVLGLDNLTDNAORAILKKHANKDKGAELKLFKAVENTLSKAAODTLKN 361  
DB 325 ftnklsehavlgldnltdnagrailkkhankdkgaalekfkavenlskaagdtlkn 384  
QY 362 AVKELTS 368  
DB 362 avkeltss 391

RESULT 3

AAB62719  
ID AAB62719 standard; Protein: 367 AA.

XX AAB62719;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 42.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PI 18-JUN-1999; 99US-0140042.

```

XX PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX DR MPI: 2001-050113/06.
XX DR N-PSDB; AAF29023.
XX PT Compositions of OspC polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PT to Lyme disease -
XX PS Claim 43; Page 102-103; 160pp; English.
XX CC The present invention provides compositions comprising ospC proteins and
XX CC chimeric ospC proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX SQ Sequence 367 AA;

Query Match 85.3%; Score 1533; DB 22; Length 367;
Best Local Similarity 86.6%; Pred. No. 1.1e-90;
Matches 318; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

OY 2 ACNNSGKDNASANSADSVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDELATKAI 61
DB 1 acnsgkdgnasansadesvkgpnlteiskitesnavlavkevellasidelatka 60
OY 62 GKRTGNGGLEANSKNTSLSGAYAISDLIAEKLNVLKNELEKIDTAKOCSTFTNKL 121
DB 61 gkrtgnggleanskntsllsgayaaisdliaeklnvlnkelelkdltakqcsftfnkl 120
OY 122 KSHAVLGIDNLTDDNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTLKNAVKEL 181
DB 121 kshavlgidnltddnagrallkxhankdkgaalekllfavenlskaadclknavekl 180
OY 182 TSPIVHGNNRSKDNASTNSADESVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDEL 241
DB 181 tsplivhgnskdgnastnsadesvkgpnlteiskitesnavlavkeletllasidel 240
OY 242 ATKAIGKKT-GNNGGLEANSKNTSLSGAYAISDLIAEKLNVLKN-BELKEKIDTAKOCS 299
DB 241 atkaigkktgngglaveagngllagayctlskltqkldglnseklikelenakcs 300
OY 300 TEFNTNLSKSHAVLGIDNLTDDNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTL 359
DB 301 tefntnlkshavlgidnltddnagrallkxhankdkgaalekllfavenlskaadclknavekl 360
OY 360 KNAVKEL 366
DB 361 ansvkel 367

RESULT 4
AAB62735
ID AAB62735 standard; Protein; 391 AA.
XX
AC AAB62735;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 74.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
PN Chimeric - Borrelia sp.
XX
PN W0200078966-A1.

```

```

XX PD 28-DEC-2000.
XX PF 19-JUN-2000; 2000MO-US16915.
XX PR 18-JUN-1999; 99US-0140042.
XX PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX DR MPI: 2001-050113/06.
XX DR N-PSDB; AAF29039.
XX PT Compositions of OspC polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PT to Lyme disease -
XX PS Claim 43; Page 142; 160pp; English.
XX CC The present invention provides compositions comprising ospC proteins and
XX CC chimeric ospC proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX SQ Sequence 391 AA;

Query Match 85.1%; Score 1530; DB 22; Length 391;
Best Local Similarity 86.4%; Pred. No. 1.8e-90;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

OY 2 ACNNSGKDNASANSADSVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDELATKAI 61
DB 25 acnsgkdgnasansadesvkgpnlteiskitesnavlavkevellasidelatka 84
OY 62 GKRTGNGGLEANSKNTSLSGAYAISDLIAEKLNVLKNELEKIDTAKOCSTFTNKL 121
DB 85 gkrtgnggleanskntsllsgayaaisdliaeklnvlnkelelkdltakqcsftfnkl 144
OY 122 KSHAVLGIDNLTDDNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTLKNAVKEL 181
DB 121 kshavlgidnltddnagrallkxhankdkgaalekllfavenlskaadclknavekl 204
OY 145 KSEHAVLGIDNLTDDNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTLKNAVKEL 204
DB 144 ksehavlgidnltddnagrallkxhankdkgaalekllfavenlskaadclknavekl 204
OY 182 TSPIVHGNNRSKDNASTNSADESVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDEL 241
DB 181 tsplivhgnskdgnastnsadesvkgpnlteiskitesnavlavkeletllasidel 264
OY 242 ATKAIGKKT-GNNGGLEANSKNTSLSGAYAISDLIAEKLNVLKN-BELKEKIDTAKOCS 299
DB 241 atkaigkktgngglaveagngllagayctlskltqkldglnseklikelenakcs 324
OY 300 TEFNTNLSKSHAVLGIDNLTDDNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTL 359
DB 301 tefntnlkshavlgidnltddnagrallkxhankdkgaalekllfavenlskaadclknavekl 384
OY 360 KNAVKEL 366
DB 385 ansvkel 391

RESULT 5
AAB62718
ID AAB62718 standard; Protein; 368 AA.
XX
AC AAB62718;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 40.
XX

```

KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 XX Chimeric - Borrelia sp.  
 PN WO20078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UNYV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PSDB; AAF29022.  
 XX  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 99-100; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. atzevili or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 368 AA;

Query Match 82.7%; Score 1485.5; DB 22; Length 368;  
 Best Local Similarity 84.3%; Pred. No. 1.2e-87;  
 Matches 311; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

OY 2 ACNNGSGKGNASANSADSVKGNLTETISKRTTESNAVLAVKEVETLLASIDELATKAI 61  
 DB 1 acnngsgkgnasansadesvkgpnlteiskrttesnavlavkevecllasidelatkai 60  
 OY 62 GKRIKNGGLEANSQKNTSLSGAVAIISDLIAEKLNVLKNELKEKIDTAKOCSTEFNTKL 121  
 DB 61 gkriknnggleansqkntslsgavaisdliaeklnvlnkeelkeidtakgscstefntkl 120  
 OY 122 KSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAODTLKNAYVEL 181  
 DB 121 ksehavlglldnltdnnaqraillkhanndkgaalekikfavenenlskaagdtlknayvel 180  
 OY 182 TSPYHGNNSRRKDGASVNSADESVKGNLTETISKRTTESNAVLAVKEVETLLASIDEL 241  
 DB 181 tsplhgnnsgkdgnstansadesvkgpnlteiskrttesnavlavavevcllasidel 240  
 OY 242 ATKAIGKRIKGN-GLEANSQKNTSLSGAVAIISDLIAEKLNVLKNELKEKIDTAKOC 299  
 DB 241 a-kaigkrikndvsidneadngslsgayaisnlttkisaisksgelkaelekakcs 299  
 OY 300 TEFTNKLKSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAO 359  
 DB 300 eeftaklkgentldgkegytdnnaakailknndktgadelkifesevknlskaakeml 359  
 OY 360 KNAVELTS 368  
 DB 360 tnsvkelts 368

RESULT 6  
 AAB62734  
 ID AAB62734 standard; Protein: 392 AA.  
 XX

AC AAB62734;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX  
 DE Borrelia sp chimeric ospC protein SEQ ID NO: 72.  
 XX  
 KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 XX Chimeric - Borrelia sp.  
 PN WO20078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UNYV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PSDB; AAF29038.  
 XX  
 XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 139-140; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. atzevili or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 392 AA;

Query Match 82.5%; Score 1482.5; DB 22; Length 392;  
 Best Local Similarity 84.0%; Pred. No. 2e-87;  
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

OY 2 ACNNGSGKGNASANSADSVKGNLTETISKRTTESNAVLAVKEVETLLASIDELATKAI 61  
 DB 25 acnngsgkgnasansadesvkgpnlteiskrttesnavlavkevecllasidelatkai 84  
 OY 62 GKRIKNGGLEANSQKNTSLSGAVAIISDLIAEKLNVLKNELKEKIDTAKOCSTEFNTKL 121  
 DB 61 gkriknnggleansqkntslsgavaisdliaeklnvlnkeelkeidtakgscstefntkl 144  
 OY 122 KSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAO 181  
 DB 121 ksehavlglldnltdnnaqraillkhanndkgaalekikfavenenlskaagdtlknayvel 180  
 OY 145 ksehavlglldnltdnnaqraillkhanndkgaalekikfavenenlskaagdtlknayvel 204  
 OY 182 TSPYHGNNSRRKDGASVNSADESVKGNLTETISKRTTESNAVLAVKEVETLLASIDEL 241  
 DB 205 tsplhgnnsgkdgnstansadesvkgpnlteiskrttesnavlavavevcllasidel 264  
 OY 242 ATKAIGKRIKGN-GLEANSQKNTSLSGAVAIISDLIAEKLNVLKNELKEKIDTAKOC 299  
 DB 265 a-kaigkrikndvsidneadngslsgayaisnlttkisaisksgelkaelekakcs 323  
 OY 300 TEFTNKLKSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAO 359  
 DB 324 eeftaklkgentldgkegytdnnaakailknndktgadelkifesevknlskaakeml 383  
 OY 360 KNAVELTS 368  
 DB 384 tnsvkelts 392

```

RESULT 7
AAB62714
ID AAB62714 standard; Protein; 369 AA.
XX
AC AAB62714;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 32.
XX
KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV ) UNIT NEW YORK STATE RES FOUND.
XX
PI (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
XX
DR N-PSDB: AAF29018.
XX
PT Compositions of ospc polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 89; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
XX
CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX
CC to Lyme disease.
XX
SQ Sequence 369 AA;

Query Match 82.3%; Score 1479; DB 22; Length 369;
Best Local Similarity 83.8%; Pred. No. 3e-87;
Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

```

```

DB 300 stettinksehavilgldntddnaqraillkhanckgaalekllkavenlskaagdt 359
QY 359 LKNAVKELETS 368
DB 360 lknavkeltss 369

RESULT 8
AAB62732
ID AAB62732 standard; Protein; 393 AA.
XX
AC AAB62732;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 68.
XX
KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV ) UNIT NEW YORK STATE RES FOUND.
XX
PI (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
XX
DR N-PSDB: AAF29036.
XX
PT Compositions of ospc polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 133-134; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
XX
CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX
CC to Lyme disease.
XX
SQ Sequence 393 AA;

Query Match 82.1%; Score 1476; DB 22; Length 393;
Best Local Similarity 83.5%; Pred. No. 5.1e-87;
Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

```





Oy	62	GK- <del>GNNGLEANOSKNTSLISGAVATSDIAEKLNVLNKNNELEKERIDPTAKOCSTERTNK</del>	120
Dd	61	gkktlqnnngldcteynhngslagayatslltkqklldglkneglknegidakkcsetftnk	120
Oy	121	LKSHAVALGLNDLTDNNAORAILKKHANKDKGALELKFRAVENTLSKAQDTLKNAVKE	180
Dd	121	lkshetldlgreyvadadakaelikltngtcktgaeelgkltseveylskaekmlansvke	180
Oy	181	LTSPIYHG-----NSNRKDGNASTNSADESVKGPNTLEISKKTITESHANVLAAYEVETLL	235
Dd	181	ltspsvaeepamvnmnsgkdgnftsasadesavkpntleiskklitesnavylvakeletll	240
Oy	236	ASIDE <del>LTKAIIGKRI-GNNGLEANSOKNTSLISGAVATSDIAEKLNVLTNK-EELER</del> ID	293
Dd	241	asidelatckalggkldqngnglweaghnngcllagayclsklltcqklldglknseklikexie	300
Oy	294	TAKOCSTERTFNKLSHAVALGLNDLTDNNAORAILKKHANKDKGALELKFRAVENTLSK	353
Dd	301	nakkcsedcfckklegeahngjdienvtdenakkailltcaakdkygaaelkfavenlak	360
Oy	354	AAQDTLKNAVKEI	366
Dd	361	aakemlansvkel	373
 RESULT 11 AAB62729 standard; Protein: 397 AA.			
XX	AAB62729;		
XX	03-APR-2001	(first entry)	
DE	Borrelia sp	chimeric ospC protein SEQ ID NO: 62.	
KM	Borrelia;	ospC; Lyme disease; vaccine; chimeric protein; tick.	
XX	Chimeric - Borrelia sp.		
OS	Chimeric - Borrelia sp.		
PN	WO200078966-A1.		
PD	28-DEC-2000.		
PF	19-JUN-2000;	2000MO-US16915.	
PR	18-JUN-1999;	99US-0140042.	
PA	(UWNY ) UNIV NEW YORK STATE RES FOUND.		
PA	(BROO-) BROOK BIOTECHNOLOGIES INC.		
PI	Dattwyler RJ, Selinost G, Dykhulzen D, Luft BJ, Gomes-Solecki M;		
DR	WPI: 2001-050113/06.		
PT	N-PsDB; AAF29033.		
XX	Compositions of OspC polypeptides from strains of Borrelia which cause		
XX	Lyme disease are used to immunize animals and detect immune responses		
XX	to Lyme disease -		
PS	Claim 43; Page 125-126; 160pp; English:		
CC	The present invention provides compositions comprising ospc proteins and		
CC	chimeric ospc proteins from members of the Borrelia genus. These may be		
CC	Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as		
CC	vaccines against Borrelia infection, which is spread by ticks and leads		
XX	to Lyme disease.		
S0	Sequence 397 AA;		

	Best Local Similarity	70.0% ; Pred. No. le-70:	
	Matches	261; Conservative	43; Mismatches
		61; Indels	8; Gaps
QY	2	ACNNSGKDGAMASADESVYKGNPLTEISKITTESNAVVLAVKEVETLLASIDELATTKAI	61
Db	25	scnnsgkdgntsansadesvgygnltelskittsdnnavllavkveallssidelaakel	84
QY	62	GKKT-GNNGLEANKNSKNTSLSGATAISDLIAEKLNVLKNELKEKIDIDAKOCSTFTPK	120
Db	85	gkktinqngldteyhngnslagayaisltlqkdldglneglkexidaakcseltfkk	144
QY	121	LKSEHAVIGLDNLTDNMQRAILTKKANKDKAAELEKIFKAVENISKAQDITLKAAVKE	180
Db	145	lkehltldgkevgtdadaeaalktngtktkpaeeigkifesevsvskaakemlansvke	204
QY	181	LTSPVHG-----NNSRKDGNASTNSADESVYKGNPLTEISKITTESNAVVLAVKEVETLL	235
Db	205	ltspvvaespanvmvmsgkdgntsansadesvgygnltelskittsdnnavllavkeltell	264
QY	236	ASIDELATTKAIKGTI-GNNGLEANKNSKNTSLSGAVALISDLIAEKLNVLKN-ELKEKID	293
Db	265	asidelatkaigkktqngnglaveevingtllagayaisltlqkdldglnsekellek	324
QY	294	TAKOCSTFTPKLKSEHAVIGLDNLTDNMQRAILTKKANKDKAAELEKTKRAVNLSEK	353
Db	325	takkscedtftkllegehaqjlenvldenaakailtdaakdkgaalekfkavenlak	384
QY	354	AAQDITLKAAVKEKEL	366
Db	385	aakemlansvkel	397

	RESULT 12
AAB62716	
ID	AAB62716 standard; Protein: 369 AA.
XX	
AC	AAB62716;
XX	
DT	03-APR-2001 (first entry)
XX	
DE	Borrelia sp chimeric ospC protein SEQ ID NO: 36.
XX	
KW	Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX	
OS	Chimeric - Borrelia sp.
OS	Chimeric - Borrelia sp.
XX	
PN	WO200078966-A1.
PD	
PD	28-DEC-2000.
PE	
PF	19-JUN-2000; 2000WO-US16915.
PR	
PR	18-JUN-1999; 99US-0140042.
PA	(UYNV ) UNIV NEW YORK STATE RES FOUND.
PA	(BROO-) BROOK BIOTECHNOLOGIES INC.
PI	Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR	WPI; 2001-050113/06.
DR	N-PSDB; AAF29020.
XX	
PT	Compositions of OspC polypeptides from strains of Borrelia which cause
PT	Lyme disease are used to immunize animals and detect immune responses
PT	to Lyme disease -
PS	
PS	Claim 43; Page 94-95; 160pp; English.
CC	
CC	The present invention provides compositions comprising ospC proteins and
CC	chimeric ospC proteins from members of the Borrelia genus. These may be
CC	Borrelia burgdorferi, B. atzelli or B. garinii. These can be used as
CC	vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 369 AA;

Query Match 67.9%; Score 1220; DB 22; Length 369;  
 Best Local Similarity 70.5%; Pred. No. 1.1e-70;  
 Matches 261; Conservative 45; Mismatches 58; Indels 6; Gaps 6;

QY 2 ACNNSGKDGNSANSADSVKGNLTETSKITTESNAVLAKEVEETLLASIDELATKAI 61  
 DB 1 acnnsygdgntsansadesvkgpnlteiskitidsnavllavkeveallssidela-kai 59  
 QY 62 GKKTGNNG-LEFANOSKNTSLSGAVALSDLAELKLVKNEE-LKEKIDTAKOCSTEEFN 119  
 DB 60 gkktkngngldneanresllagaytstltqklsklingseggkelaakcseefst 119  
 QY 120 KLKSEHVLGIDNLTDDNAORAILKKH-ANKDKGALEKLFKAVENLSKAADTLKNAY 178  
 DB 120 klkshbgqlgqyvdnakaalkanaagkdkyveeleklsgslselskaakemiansv 179  
 QY 179 KELTSPYVHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVLAKEVEETLLAS 238  
 DB 180 keltspvvhgnsygdgntsansadesvkgpnlteiskitidsnavllavkeveetllasi 239  
 QY 239 DELATTAIGKRT-GNNGLEFANOSKNTSLSGAVALSDLAELKLVKN-EELKEKIDTAK 296  
 DB 240 delatkaigkrtkngngllaveagngtllagaytstltqkldgklnseklkelenak 299  
 QY 297 QCSTEEFNKLSSEHVALGIDNLTDDNAORAILKKHANKDKGALEKLFKAVENLSKAQ 356  
 DB 300 kcsedftkklgeghaqglgenvdenakallttdaakdgaaeleklfkavenlakaak 359  
 QY 357 DTAKNAVVEL 366  
 DB 360 emlansvkel 369

RESULT 13

AAB62739

ID AAB62739 standard; Protein; 400 AA.

AC AAB62739;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UANY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29043.

Compositions of ospc polypeptides from strains of Borrelia which cause  
 Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -

PS Claim 43; Page 153; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

SQ Sequence 400 AA;

Query Match 67.9%; Score 1219.5; DB 22; Length 400;  
 Best Local Similarity 69.4%; Pred. No. 1.3e-70;  
 Matches 261; Conservative 43; Mismatches 61; Indels 11; Gaps 4;

QY 2 ACNNSGKDGNSANSADSVKGNLTETSKITTESNAVLAKEVEETLLASIDELATKAI 61  
 DB 25 acnnsygdgntsansadesvkgpnlteiskitidsnavllavkeveallssidela-kai 84  
 QY 62 GKKTGNNG-LEFANOSKNTSLSGAVALSDLAELKLVKNEE-LKEKIDTAKOCSTEEFN 120  
 DB 85 gkktkngngldneanresllagaytstltqklsklingseggkelaakcseefst 144  
 QY 121 KLKSEHVLGIDNLTDDNAORAILKKHANKDKGALEKLFKAVENLSKAADTLKNAY 180  
 DB 145 keltspvvhgnsygdgntsansadesvkgpnlteiskitidsnavllavkeveetllasi 204  
 QY 181 LTSPYVHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVLAKEVEE 232  
 DB 205 ltsprvaesprkpsvvhgnsygdgntsansadesvkgpnlteiskitidsnavllavkele 264  
 QY 233 TLASIDELATKAIGKRT-GNNGLEFANOSKNTSLSGAVALSDLAELKLVKN-EELKE 290  
 DB 265 tlasidelatkaigkrtkngngllaveagngtllagaytstltqkldgklnseklkelenak 324  
 QY 291 KIDTAKOCSTEEFNKLSSEHVALGIDNLTDDNAORAILKKHANKDKGALEKLFKAVEN 350  
 DB 325 klenakkcsedftkklgeghaqglgenvdenakallttdaakdgaaeleklfkaven 384  
 QY 351 LSKAADTLKNAYVEL 366  
 DB 385 lakaakemlansvkel 400

RESULT 14

AAB62731

ID AAB62731 standard; Protein; 393 AA.

AC AAB62731;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 66.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UANY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:34 ; Search time 55.5 Seconds

(without alignments)  
149.211 Million cell updates/sec

Title: US-09-596-746A-38

Perfect score: 1797

Sequence: 1 MACNNSGKDDGNASANSADSE.....ENLSKAQDTLKNVKELTSS 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

212252 seqs, 22503292 residues

212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	37.5	466	4	US-08-235-836C-110 Sequence 110, App
2	657.5	36.6	212	1	US-08-158-353-2 Sequence 2, App1
3	640	35.6	588	4	US-08-235-836C-122 Sequence 122, App
4	638.5	35.5	466	4	US-08-235-836C-107 Sequence 107, App
5	623.5	34.7	209	4	US-09-196-293-15 Sequence 15, App1
6	623.5	34.7	210	1	US-08-158-353-3 Sequence 3, App1
7	617.5	34.4	210	4	US-08-235-836C-30 Sequence 30, App1
8	616.5	34.3	210	4	US-08-209-603E-15 Sequence 15, App1
9	588.5	32.7	212	1	US-08-158-353-4 Sequence 4, App1
10	584.5	32.5	212	1	US-08-031-295-2 Sequence 2, App1
11	584.5	32.5	212	4	US-07-903-580-2 Sequence 2, App1
12	583.5	32.5	212	4	US-08-235-836C-34 Sequence 34, App1
13	578.5	32.2	212	4	US-09-196-293-11 Sequence 11, App1
14	578.5	32.2	212	4	US-08-209-603E-11 Sequence 11, App1
15	563	31.3	207	4	US-08-235-836C-36 Sequence 36, App1
16	536	29.8	209	4	US-08-235-836C-32 Sequence 32, App1
17	157.5	8.8	1588	5	PCT-US93-07261-11 Sequence 11, App1
18	157.5	8.8	1663	5	PCT-US93-07261-16 Sequence 16, App1
19	150	8.3	695	1	US-08-127-499A-23 Sequence 23, App1
20	150	8.3	695	1	US-08-482-847-23 Sequence 23, App1
21	150	8.3	1147	3	US-08-470-260-5 Sequence 5, App1
22	150	8.3	1147	3	US-08-471-491-5 Sequence 5, App1
23	150	8.3	1147	4	US-08-466-662-5 Sequence 5, App1
24	150	8.3	3289	2	US-08-477-451-2 Sequence 2, App1
25	149.5	8.3	1196	1	US-08-144-121-4 Sequence 4, App1
26	149.5	8.3	1196	2	US-08-735-893-4 Sequence 4, App1
27	149	8.3	630	4	US-08-973-462-9 Sequence 9, App1

28	141.5	7.9	1786	4	US-08-973-462-8 Sequence 8, App1
29	141	7.8	859	1	US-08-053-614-2 Sequence 2, App1
30	141	7.8	859	1	US-08-316-397B-2 Sequence 2, App1
31	141	7.8	859	2	US-09-034-306-2 Sequence 2, App1
32	141	7.8	859	4	US-09-259-437-2 Sequence 2, App1
33	141	7.8	859	5	PCT-US93-09782-2 Sequence 2, App1
34	141	7.8	1181	1	US-08-053-614-4 Sequence 4, App1
35	141	7.8	1181	1	US-08-316-397B-4 Sequence 4, App1
36	141	7.8	1181	2	US-09-034-306-4 Sequence 4, App1
37	141	7.8	1181	5	PCT-US93-09782-4 Sequence 4, App1
38	141	7.8	1181	5	PCT-US93-09782-4 Sequence 4, App1
39	139.5	7.7	2482	1	US-08-328-254-6 Sequence 6, App1
40	138.5	7.7	619	1	US-08-465-746-2 Sequence 2, App1
41	138.5	7.7	619	1	US-08-214-164-2 Sequence 2, App1
42	138.5	7.7	619	2	US-08-467-852A-3 Sequence 3, App1
43	138.5	7.7	619	2	US-08-246-636-2 Sequence 2, App1
44	138.5	7.7	619	2	US-08-247-491A-3 Sequence 3, App1
45	138.5	7.7	619	2	US-08-319-795-2 Sequence 2, App1

#### ALIGNMENTS

RESULT 1  
US-08-235-836C-110  
Sequence 110, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562zel Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Thereof  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESS: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110  
Query Match 37.5%; Score 673; DB 4; Length 466;  
Best Local Similarity 44.9%; Pred. No. 1.3e-45;  
Matches 173; Conservative 58; Mismatches 112; Indels 42; Gaps 11;  
OY 1 MACNNSGKDDGNASANSADSEYKGNPLTEISKKTIESNNAVYLAIVEVEFLASIDELATKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNLTETISKTTIDSNVAVLAVKEVETLLASIDELATAKA 76  
QY 61 IGGKI-GANGLEANOSKNTSLSGAVATSDIAEKLNVAKNELEKIDTAKOCSTETN 119  
Db 77 IGGKIHNNGNDJTEYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKOCSEFTN 136  
QY 120 KIKSEHAVIGLNDLTJDNQRAILKHKANKGAELKFKAVENLSKAQDTLKNAVK 179  
Db 137 KIKAKHHDLDGEGYTDADAKAIIKTNGTGTGAEEGLKLFESVEVLAKAKMELANSVK 196  
QY 180 ELTSPPIVHGNSRRKGNAST---NSADESVKGNLTETISKTTESNAVVLAVKEVETL 235  
Db 197 ELTSPVVAESPKRKQVSSLEKNSVSDLPGBMKVLVSKENKQCKY-----DLI 248  
QY 236 ASIDELATKAIKGIKGN-----GLEANSKNTSLSGAVATSDIAEKLNVAKNEE-- 287  
Db 249 ATYDKLEKLTSDK--NNGSGVLEGVAKDSKVKLTIS-----DDIGQTLLEVFKEKEDGKT 301  
QY 288 -LKEKIDTAKOCSTE--FTNKLKSEHAVIGLNDLTJDNQRAILKHKANKGAELK 344  
Db 302 LVSKKVTSKRKSSTEEKFNEKEVSEKII-----TRADGTLEETGTGKSDSGKAKEY 354  
QY 345 FK--AVENLSKAQDTLKNAVKELT 367  
Db 355 LKGYVLEGTILTAERTTL--VVKEGT 377

RESULT 2  
US-08-158-353-2  
; Sequence 2, Application US/08158353  
; Patent No. 5620862  
; GENERAL INFORMATION:  
; APPLICANT: Padula, Steven J.  
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
; TITLE OF INVENTION: Disease  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Millita Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/158,353  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Alice O.  
; REGISTRATION NUMBER: 33,542  
; REFERENCE/DOCKET NUMBER: UCT93-05  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 212 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-158-353-2

Query Match 36.6%; Score 657.5; DB 1; Length 212;  
Best Local Similarity 72.9%; Pred. No. 7,6e-45;  
Matches 140; Conservative 20; Mismatches 29; Indels 3; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETISKTTESNAVVLAVKEVETLLASIDELATAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNLTETISKTTESNAVVLAVKEVETLLASIDELATAKA 76  
QY 61 IGGKIHNNGNDJTEYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKOCSEFTN 119  
Db 77 IGGKIHNNGNDJTEYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKOCSEFTN 136  
QY 120 KIKSEHAVIGLNDLTJDNQRAILKHKANKGAELKFKAVENLSKAQDTLKNAVK 177  
Db 137 KIKSSAEELGINGAATDANAKAAILKTNGTGTGAQELKLFESVKMLSKRAQDTLLNS 196  
QY 178 VKELTSPVHGN 189  
Db 197 VKELTSPVVAEN 208

RESULT 3  
US-08-235-836C-122  
; Sequence 122, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Luft, Benjamin J.  
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235,836C  
; FILING DATE: 29-Apr-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; REFERENCE/DOCKET NUMBER: BNI93-28A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729  
; INFORMATION FOR SEQ ID NO: 122:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 588 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-836C-122

Query Match 35.6%; Score 640; DB 4; Length 588;  
Best Local Similarity 39.1%; Pred. No. 6,8e-43;  
Matches 184; Conservative 55; Mismatches 110; Indels 122; Gaps 17;

QY 7 GKDGNANSADSVKGN-----PNLTETISKTTESNAVVLAVKEVET---LTAASID 54  
Db 9 GSGKENDLNLEDSKSHONAKODLPAYTEVSLSFNGKKIFVSEKNSSGCYDILRAITD 68  
QY 55 ELATKAIKGIKINNG--LEANO-SKNTSLSGAVATSDIAEKLNVAKNELEKIDTAK 110

```

Db 69 QVELKGTSDK--NNGSGTLEGSKPKSKYKLVTSADLNTVLEAFDA-SNQKISSKV-TK 124
Qy 111 KQCS-TEFT---NKL-----KSEHAVLGIDNLT-DNORAI-----142
Db 125 KQSTTEETLKAKKLSKRLTRNSGTTLEYSQITDADNATKAVETLKNSIKLEGLVVGK 184
Qy 143 -----LKKHANKD-----150
Db 185 TTYEIEGCVTLTKREIEKDGKVVFLNDTPAGSKKTKGKWDSTLTLSADSKTKDVLV 244
Qy 151 -----KGAELEKLEKFAVENLSKRAODTLKNAVKELSPVTHGNSRKDNAS 198
Db 245 LTDCGTTVOOYNAGTSLESASAEIKNLSE-----LKNALK--GHPH--GNNGKQNTS 295
Qy 199 TNSADESVKGNLTETISKTTESNAVYLAKEVETLLASIDELATKAIKKTI-GNNGLEA 257
Db 296 ANSADESVKGNLTETISKTTESNAVYLAKEVEALLSIDELATKAIKKTIHONNGIDT 355
Qy 258 NOSKNTSLSGAVALDILAEKLVNLEKNEELKEIDTAKCSTETPKKSEHAVLGIDN 317
Db 356 EYNHNSLLAGRYAISTLIKOKLDGKNEGLKEKIDAKKCSFTPKKLEKHTDLGKEG 415
Qy 318 LTDMNORAILKKHANKDGALEKLEKFAVENLSKRAODTLKNAVKELTS 368
Db 416 VTDADKAEALTKNGTKTGAEELGKLFESVEVLSNAKAKMLANSVKELTS 466

RESULT 4
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-107
Query Match 35.5%; Score 638.5; DB 4; Length 466;

```

```

Best Local Similarity 44.6%; Pred. No. 6,6e-43;
Matches 166; Conservative 50; Mismatches 119; Indels 37; Gaps 8;

Qy 19 ESKVGNLTETISKTTESNAVYLAKEVEETLLASIDELATKAIKKTI-GNNGLEANS-KN 77
Db 100 EFKEDKGTLSKVTSKDSSTEEKFNEKEVS-EKITRADGTRELYGDKSGSKA 158
Qy 78 TSLSGAVALSDILAEKLVNLEKNEELKEIDTAKCSTETPKKSEHAVLGIDNLT 134
Db 159 KEVLKGVLEGTTLAEKTLVKEGTVTLKSNLSKSGSEVELNDTSS--AATKTRAW 216
Qy 135 DDNAQRILKKHANKD-----KGAELEKLEKFAVENLSKRAODTLKNA 177
Db 217 NDSTLTISADSKTKTDVFLFTGTTTVOOYNAGTSLESASAEIKNLSE-----LKNA 271
Qy 178 VKELTSPVTHGNSRKDNASTNSADESVKGNLTETISKTTESNAVYLAKEVEETLLAS 237
Db 272 LK-----MANNKGDKNTSANSADSVKGNLTETISKTTESNAVYLAKEVEALLSS 324
Qy 238 IDELATKAIKKTI-GNNGLEANSKNTSLSGAVALDILAEKLVNLEKNEELKEIDTAK 296
Db 325 IDETAAKAIKKTIHONNGIDTVEYNHNSLLAGRYAISTLIKOKLDGKNEGLKEKIDAK 384
Qy 297 OCSTETPKKSEHAVLGIDNLTDMNORAILKKHANKDGALEKLEKFAVENLSKRAO 356
Db 385 KCSETFTNKLAKHTDGLKEGVTADAKAEALTKNGTKTGAEELGKLFESVEVLSNAK 444
Qy 357 DTLKNAVKELTS 368
Db 445 EMLANSVKELTS 456

RESULT 5
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Moltz, Manfred
; APPLICANT: Soutscheck, Erwin
; APPLICANT: Wilske, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; FILE REFERENCE: 738.001052
; CURRENT APPLICATION NUMBER: US/09/196,293
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1988-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-196-293-15
Query Match 34.7%; Score 623.5; DB 4; Length 209;
Best Local Similarity 69.3%; Pred. No. 3.5e-42;
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;
Qy 1 MACNNSGKDNASANSADSVKGNLTETISKTTESNAVYLAKEVEETLLASIDELATKA 60
Db 17 ISCNNSGKDNSTANSADSVKGNLTETISKTTESNAVYLAKEVEALLSIDELAKA 76

```





```

Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFENER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 07/862,535
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3335
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match          34.3%, Score 616.5; DB 4; Length 210;
Best Local Similarity 69.0%, Pred. No. 1.3e+41;
Matches 129; Conservative 24; Mismatches 33; Indels 1; Gaps 1

Oy      1 MACNNSGKDNAGNASADSVKGPVLTEISKRTESNAVAVLVKVEETLASIDELETAKA 60
        : ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      17 ISCNNSGDKGNTSNASADESVKGPLTEISKRTITSMNAVLLVLEVALLSIDELAAKA 76
        : ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Oy      61 IGKKI-GNGGLEANSKNTSLSGAVAISDLIAEKLVNLNELEKIDIPAROCSTFTFN 119
        : ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      77 IGRKIHNNGNDLPTENNHHNSLLAGAVAI STLTKRKLDGLKNGLKRIIDAKKCSFTFTN 136
        : ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

120 KLSKHAVLGIDLNTDDNAQRALIKKHKANKDGAALELEKFAVAENVLSAAODTLKNAVY 179
        : ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Db    137 KLRKHTDIAKEGVYTDADAKFAILKTNGTKRKGAELGKLPEESVYLSKAAKEMLANSVK 196
      180 ELTSPIV 186
           |||||:1
Db    197 ELTSPV 203

RESULT   9
US-08-158-353-4
; Sequence 4, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-6240
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match          32.7%; Score 588.5; DB 1; Length 212;
Best Local Similarity 68.8%; Pred. No. 2,1e-39;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3,

OY     1 MACNNSGKD-G-NASANSADESVKGNTLIEISKITTESNAVNVLAVKEVEETILLASIDELATK 59
      :|||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    17 ISCNNSGGKGSASPNDPDESAGKNLTETISKRIDSNAFVAIVEVEETILLASIDELAK 76
      :|||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY     60 AIGKRI -GNNGSEANOSKNTSLSGAVAIVSDIJAELKLVLN -EELKERIDPAKOCSTEF 117
      |||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    77 AIGQITDNNNGIALANNONGSLLACAYAI STLTERKSLRKLHBEIKRTIAAKKCSPEF 136
      |||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY    118 TNKLKSEHAVIGLDNI.TDDMAQRALILKKRANKDKGAEELEKLFKAVENLSKAAODTLKNA 177
      |||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    137 TNKLKSAGHDIGKOQDATDHDHAKAAILKTHATVDKGAKEPKDFESVEGLLKAAQVALTWTS 196
      |||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY     178 VKELTSPIV 186
           |||||:1
Db    197 VKELTSPV 203

```

Db 197 VKELSPV 205

RESULT 11  
US-07-903,580-2  
; Sequence 2, Application US/07903580  
; Patent NO. 6221363

RESULT 12  
US-08-235-836C-34  
; Sequence 34, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Lafty, Benjamin J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-34

Query Match 32.5%; Score 583.5; DB 4; Length 212;  
Best Local Similarity 68.3%; Pred. No. 5.1e-39;  
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNNGKDC-MASANSADSVKGNLTETSKRTTESNAVAVLVKVEETLASELATAK 59  
DB 17 ISCSNKGKGDSTNPADBSAGPNLTETSKRTTDSNAVLAVKEETVLISDELAKK 76  
60 AIGKKI-GNNGLEANSKNTSLGAYAISDLIAEKLNVLKN-BELKEKIDTAKOCSTEF 117  
DB 77 AIGKIDNNNGGLAANNQNSLAGAYAI STLITERKLSKLNLELKTETAKKCKCEE 136  
QY 118 TNKLKSEHAVLGIDNLTDDNAQRALIKKHANKDKGAEELEKLEKRAVENLSKAADTLKNA 177  
DB 137 TNKLKSGHADLGKODATDDHAKAAIILKTHATTDGKAKEFKDLFESVGLLKAQVALTNS 196  
QY 178 VKELTSPIV 186  
DB 197 VKELTSPV 205

RESULT 13  
US-09-196-293-11  
Sequence 11, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Sautscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: 738.001US2  
CURRENT APPLICATION NUMBER: US/09/196,293

CURRENT FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 32.2%; Score 578.5; DB 4; Length 212;  
Best Local Similarity 68.3%; Pred. No. 1.3e-38;  
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

QY 1 MACNNGKDC-MASANSADSVKGNLTETSKRTTESNAVAVLVKVEETLASELATAK 59  
DB 17 ISCSNKGKGVILSTNPADBSAGPNLTETSKRTTDSNAVLAVKEETVLISDELAKK 76  
QY 60 AIGKKI-GNNGLEANSKNTSLGAYAISDLIAEKLNVLKN-BELKEKIDTAKOCSTEF 117  
DB 77 AIGKIDNNNGGLAANNQNSLAGAYAI STLITERKLSKLNLELKTETAKKCKCEE 136  
QY 118 TNKLKSEHAVLGIDNLTDDNAQRALIKKHANKDKGAEELEKLEKRAVENLSKAADTLKNA 177  
DB 137 TNKLKSGHADLGKODATDDHAKAAIILKTHATTDGKAKEFKDLFESVGLLKAQVALTNS 196  
QY 178 VKELTSPIV 186  
DB 197 VKELTSPV 205

RESULT 14  
US-08-209-603E-11  
Sequence 11, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILSKE, BETTINA  
APPLICANT: PREAC-MURISIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOOTSCHHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:45 ; Search time 621.2 Seconds  
(without alignments)  
164.485 Million cell updates/sec

Title: US-09-596-746A-38

Perfect score: 1797  
Sequence: 1 MACNNSGKDDGNASANSADSE.....ENLSKAQDTLKNAVKELTS 368

## Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1797	100.0	368	19	US-09-596-746-38
2	1797	100.0	368	19	US-09-596-746A-38
3	1789	99.6	391	19	US-09-596-746-76
4	1789	99.6	392	19	US-09-596-746A-76
5	1538	85.6	368	19	US-09-596-746A-42
6	1533	85.3	367	19	US-09-596-746-42
7	1530	85.1	391	19	US-09-596-746-74
8	1530	85.1	392	19	US-09-596-746A-74
9	1490.5	82.9	369	19	US-09-596-746A-40

10	1485.5	82.7	368	19	US-09-596-746-40	Sequence 40, Appl
11	1484	82.6	370	19	US-09-596-746A-32	Sequence 32, Appl
12	1482.5	82.5	392	19	US-09-596-746-72	Sequence 72, Appl
13	1482.5	82.5	393	19	US-09-596-746A-72	Sequence 72, Appl
14	1479	82.3	369	19	US-09-596-746-32	Sequence 32, Appl
15	1476	82.1	393	19	US-09-596-746-68	Sequence 68, Appl
16	1476	82.1	394	19	US-09-596-746A-68	Sequence 68, Appl
17	1429	68.4	374	19	US-09-596-746A-76	Sequence 26, Appl
18	1429	68.4	377	19	US-09-596-746A-76	Sequence 26, Appl
19	1225.5	68.2	377	19	US-09-596-746-30	Sequence 30, Appl
20	1225	68.2	370	19	US-09-596-746A-36	Sequence 36, Appl
21	1224	68.1	373	19	US-09-596-746-26	Sequence 26, Appl
22	1221	67.9	397	19	US-09-596-746-62	Sequence 62, Appl
23	1221	67.9	398	19	US-09-596-746A-62	Sequence 62, Appl
24	1220	67.9	369	19	US-09-596-746-36	Sequence 36, Appl
25	1219.5	67.9	400	19	US-09-596-746-82	Sequence 82, Appl
26	1219.5	67.9	401	19	US-09-596-746A-82	Sequence 82, Appl
27	1217	67.7	393	19	US-09-596-746-66	Sequence 66, Appl
28	1217	67.7	394	19	US-09-596-746A-66	Sequence 66, Appl
29	1181.5	65.7	375	19	US-09-596-746A-24	Sequence 24, Appl
30	1178	65.6	378	19	US-09-596-746A-28	Sequence 28, Appl
31	1178	65.6	378	19	US-09-596-746A-28	Sequence 28, Appl
32	1177.5	65.5	371	19	US-09-596-746A-34	Sequence 34, Appl
33	1176.5	65.5	374	19	US-09-596-746A-24	Sequence 24, Appl
34	1173.5	65.3	398	19	US-09-596-746-60	Sequence 60, Appl
35	1173.5	65.3	399	19	US-09-596-746A-60	Sequence 60, Appl
36	1172.5	65.2	370	19	US-09-596-746-34	Sequence 34, Appl
37	1172	65.2	401	19	US-09-596-746-80	Sequence 80, Appl
38	1172	65.2	402	19	US-09-596-746A-80	Sequence 80, Appl
39	1169.5	65.1	394	19	US-09-596-746-64	Sequence 64, Appl
40	1169.5	65.1	395	19	US-09-596-746A-64	Sequence 64, Appl
41	1167.5	65.0	560	19	US-09-596-746A-52	Sequence 52, Appl
42	1167.5	65.0	560	19	US-09-596-746A-52	Sequence 52, Appl
43	1162.5	64.7	385	19	US-09-596-746A-56	Sequence 56, Appl
44	1157.5	64.4	384	19	US-09-596-746-56	Sequence 56, Appl
45	1154.5	64.2	408	19	US-09-596-746-78	Sequence 78, Appl

## ALIGNMENTS

RESULT 1  
US-09-596-746-38  
Sequence 38, Application US/09596746  
GENERAL INFORMATION:  
APPLICANT: Datwyler, Raymond J.  
APPLICANT: Seino, Gerald  
APPLICANT: Dykhuizen, Daniel  
APPLICANT: Luft, Benjamin J.  
APPLICANT: Maria J.C. Gomes-Solecki  
TITLE OF INVENTION: Groups of Botrelia burgdorferi and Botrelia azeiella That Cause Lyme Disease In Humans  
FILE REFERENCE: 2631.1002-001  
CURRENT APPLICATION NUMBER: US/09/596, 746  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/140,042  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 38  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: OSCP Chimera  
US-09-596-746-38  
Query Match 100.0%; Score 1797; DB 19; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.4e-124;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```

RESULT 4
US-09-596-746A-76
; Sequence 76, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhuiszen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; EO ID NO 76
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-76

Query Match          99.6%; Score 1789; DB 19; Length 392;
Best Local Similarity 99.7%; Pred. No. 6.1e-124;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNSGKDGNASANSADSEYKGNLTETISKITESNAVYLAVKEVETLLASIDELATKAI 61
DB 26 SCNSGKDGNASANSADSEYKGNLTETISKITESNAVYLAVKEVETLLASIDELATKAI 85
QY 62 GKRIKNGGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEFNTKL 121
DB 86 GKRIKNGGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEFNTKL 145
QY 122 KSHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKNVKE 181
DB 146 KSHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKNVKE 205
QY 182 TSPVHGNNSRKDGNASTNSADESVKGNLTETISKITESNAVYLAVKEVETLLASIDEL 241
DB 206 TSPVHGNNSRKDGNASTNSADESVKGNLTETISKITESNAVYLAVKEVETLLASIDEL 265
QY 242 ATRAIGKIKGNNGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEF 301
DB 266 ATRAIGKIKGNNGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEF 325
QY 302 FTKKLSEHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKN 361
DB 326 FTKKLSEHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKN 385
QY 362 AVKEKLS 368
DB 386 AVKEKLS 392

```

```

RESULT 5
US-09-596-746A-42
; Sequence 42, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhuiszen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; EO ID NO 42
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
US-09-596-746A-42

Query Match          85.6%; Score 1538; DB 19; Length 368;
Best Local Similarity 86.7%; Pred. No. 2.1e-105;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

QY 1 MACNSGKDGNASANSADSEYKGNLTETISKITESNAVYLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGNASANSADSEYKGNLTETISKITESNAVYLAVKEVETLLASIDELATKA 60
QY 61 ICKRIKNGGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEFNTKL 120
DB 61 ICKRIKNGGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEFNTKL 120
QY 121 KSHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKNVKE 180
DB 121 KSHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDTLKNVKE 180
QY 181 LTSPVHGNNSRKDGNASTNSADESVKGNLTETISKITESNAVYLAVKEVETLLASIDE 240
DB 181 LTSPVHGNNSRKDGNASTNSADESVKGNLTETISKITESNAVYLAVKEVETLLASIDE 240
QY 241 LATRAIKKTI-GNNGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEF 298
DB 241 LATRAIKKTI-GNNGLEANSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQCSTEF 300
QY 299 STEFTKLSSEHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDT 358
DB 301 SEDFTKLSSEHAVLGLDNLTDNAQRILKHKHNDKGAEELEKLFRAVENLSKAADDT 360
QY 359 LKNVKEK 366
DB 361 LANSVKEK 368

```

Best Local Similarity 86.6%; Pred. No. 4.9e-105;  
Matches 317; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
DB 1 ACNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 60
OY 62 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 121
DB 61 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 120
OY 122 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 181
DB 121 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 180
OY 182 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
DB 181 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 240
OY 242 ATKAIGKRI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKN-EELKEKIDTAQOC 299
DB 241 ATKAIGKRIQOONGGLAVEAGHNGTLLAGAYTISKLTOKLDGKNSSEKLEKENNAKKS 300
OY 300 TEFTNKLKSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTL 359
DB 301 EFTTKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKFKAVENLSKAQDTL 360
OY 360 KNAVAVKEL 366
DB 361 ANSVKEL 367

RESULT 7
US-09-596-746-74
; Sequence 74, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 391
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-74
```

Query Match 85.1%; Score 1530; DB 19; Length 391;  
Best Local Similarity 86.4%; Pred. No. 8.9e-105;

Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
DB 25 SCNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 84
OY 62 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 121
DB 85 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 144
OY 122 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 181
DB 145 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 204
```

```
OY 182 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
DB 205 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 264
OY 242 ATKAIGKRI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKN-EELKEKIDTAQOC 299
DB 265 ATKAIGKRIQOONGGLAVEAGHNGTLLAGAYTISKLTOKLDGKNSSEKLEKENNAKKS 324
OY 300 TEFTNKLKSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTL 359
DB 325 EFTTKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKFKAVENLSKAQDTL 384
OY 360 KNAVAVKEL 366
DB 385 ANSVKEL 391
```

```
RESULT 8
US-09-596-746A-74
; Sequence 74, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-74
```

Query Match 85.1%; Score 1530; DB 19; Length 392;  
Best Local Similarity 86.4%; Pred. No. 8.9e-105;  
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
DB 26 SCNNSGKDNASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 85
OY 62 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 121
DB 86 GKRIKNGNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTEFNTKL 145
OY 122 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 181
DB 146 KSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTLKNVAVKEL 205
OY 182 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
DB 206 TSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 265
OY 242 ATKAIGKRI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKN-EELKEKIDTAQOC 299
DB 266 ATKAIGKRIQOONGGLAVEAGHNGTLLAGAYTISKLTOKLDGKNSSEKLEKENNAKKS 325
OY 300 TEFTNKLKSEHAVLIGDNLTDNAQRAILKKHANKDGAELKFKAVENLSKAQDTL 359
DB 326 EFTTKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKFKAVENLSKAQDTL 385
OY 360 KNAVAVKEL 366
DB 386 ANSVKEL 392
```

```
RESULT 9
US-09-596-746A-40
; Sequence 40, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-40

Query Match      82.9%; Score 1490.5; DB 19; Length 369;
Best Local Similarity 84.3%; Pred. No. 6.9e-102;
Matches 312; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

QY 1 MACNSGKDGNAASNADESVMKPNLTKTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGNAASNADESVMKPNLTKTESNAVAVLAVKEVETLLASIDELATKA 60
QY 61 IGGKIGNGLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOCSTEFNTK 120
DB 61 IGGKIGNGLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOCSTEFNTK 120
QY 121 LKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 180
DB 121 LKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 180
QY 121 LKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 180
DB 121 LKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 180
QY 181 LSPYHGNNSKRDGNASTNSADESVKPNLTKTESNAVAVLAVKEVETLLASIDE 240
DB 181 LSPYHGNNSKRDGNASTNSADESVKPNLTKTESNAVAVLAVKEVETLLASIDE 240
QY 241 LATRAIGKTKIGNN-GLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOC 298
DB 241 LA-KAIGKTKIGNN-GLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOC 298
QY 299 STEFTNKLKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOT 358
DB 299 STEFTNKLKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOT 358
QY 359 LKNAVKEKLT 368
DB 359 LKNAVKEKLT 368
QY 360 LTNVSKELTS 369
DB 360 LTNVSKELTS 369

RESULT 10
US-09-596-746-40
; Sequence 40, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
```

```
FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-40

Query Match      82.7%; Score 1485.5; DB 19; Length 368;
Best Local Similarity 84.3%; Pred. No. 1.6e-101;
Matches 311; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNNSGKDGNAASNADESVMKPNLTKTESNAVAVLAVKEVETLLASIDELATKAI 61
DB 2 ACNNSGKDGNAASNADESVMKPNLTKTESNAVAVLAVKEVETLLASIDELATKAI 61
QY 62 GKKIGNGLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOCSTEFNTK 121
DB 62 GKKIGNGLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOCSTEFNTK 121
QY 122 KSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 181
DB 122 KSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTLKNAVKE 181
QY 182 TSPIVHGNNSKRDGNASTNSADESVKPNLTKTESNAVAVLAVKEVETLLASIDE 241
DB 182 TSPIVHGNNSKRDGNASTNSADESVKPNLTKTESNAVAVLAVKEVETLLASIDE 241
QY 242 ATRAIGKTKIGNN-GLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOC 299
DB 242 A-KAIGKTKIGNN-GLEANOSKNTSLSGAIVASDLAEKLVNKLNELEKIDTAKOC 299
QY 300 TEFTNKLKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTL 359
DB 300 TEFTNKLKSEHAVLGDLNLTDDNAORAILKKHANKDKGAELKLFKAVENLSKAADOTL 359
QY 360 KNAVKEKLT 368
DB 360 KNAVKEKLT 368
QY 360 LTNVSKELTS 368
DB 360 LTNVSKELTS 368

RESULT 11
US-09-596-746A-32
; Sequence 32, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
```

US-09-596-746a-32

Query Match 82.6%; Score 1484; DB 19; Length 370;  
 Best Local Similarity 83.8%; Pred. No. 2.1e-101;  
 Matches 311; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

QY 1 MACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 60  
 DB 1 MACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 59  
 QY 61 IGKRTGNNG-LEANSKNTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOCSTEF 118  
 DB 60 IGKRTGNNGSLDNEANRRESLLAGAVITSTLTQKLSKNGSEGLKEKIAAKKCESEFS 119  
 QY 119 NKLSEHAVIGLIDNTDNOARAILKHK-ANKDKGAELFKFAVENLSKAODTLKNA 177  
 DB 120 TKLKNHAGLQIGVTDENAKKAILKNAACKDKGVELEKLSLSSLSKAAKEMLANS 179  
 QY 178 VKELSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLAS 237  
 DB 180 VKELSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLAS 239  
 QY 238 IDELATKAIGKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKO 297  
 DB 240 IDELATKAIGKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKO 299  
 QY 298 CSTFTNKLKSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAOD 357  
 DB 300 CSTFTNKLKSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAOD 359  
 QY 358 TLKNAVELTS 368  
 DB 360 TLKNAVELTS 370

RESULT 12  
 US-09-596-746-72  
 ; Sequence 72, Application US/09596746  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Datwyler, Raymond J.  
 ; APPLICANT: Sehnost, Gerald  
 ; APPLICANT: Dykhuitzen, Daniel  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746  
 ; PRIOR FILING DATE: 2000-06-16  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 72  
 ; LENGTH: 392  
 ; TYPE: PRT  
 ; ORGANISM: ospC Chimera  
 US-09-596-746-72

Query Match 82.5%; Score 1482.5; DB 19; Length 392;  
 Best Local Similarity 84.0%; Pred. No. 2.9e-101;  
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 61  
 DB 25 SCNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 84  
 QY 62 GKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTEF 121  
 DB 85 GKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTEF 144

QY 122 KSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTLKNAVKEL 181  
 DB 145 KSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTLKNAVKEL 204  
 QY 182 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 241  
 DB 205 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 264  
 QY 242 ATKAIGKRTGNNG-LEANSKNTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 299  
 DB 265 A-KAIGKRTGNNGSLDNEADHNGSLISGAVITSTLTQKLSKNGSEGLKEKIAAKKCESEFS 323  
 QY 300 TEFNTNKLSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTL 359  
 DB 324 TEFNTNKLSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTL 383  
 QY 360 KNAVKELTS 368  
 DB 384 TNSVKELTS 392

RESULT 13  
 US-09-596-746a-72  
 ; Sequence 72, Application US/09596746a  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Datwyler, Raymond J.  
 ; APPLICANT: Sehnost, Gerald  
 ; APPLICANT: Dykhuitzen, Daniel  
 ; APPLICANT: Luft, Benjamin J.  
 ; APPLICANT: Maria J.C. Gomes-Solecki  
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 ; FILE REFERENCE: 2631.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/596,746a  
 ; PRIOR FILING DATE: 2000-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/140,042  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 72  
 ; LENGTH: 393  
 ; TYPE: PRT  
 ; ORGANISM: ospC Chimera  
 US-09-596-746a-72

Query Match 82.5%; Score 1482.5; DB 19; Length 393;  
 Best Local Similarity 84.0%; Pred. No. 2.9e-101;  
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 61  
 DB 26 SCNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 85  
 QY 62 GKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTEF 121  
 DB 86 GKRTGNNGLEANSKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTEF 145  
 QY 122 KSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTLKNAVKEL 181  
 DB 146 KSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTLKNAVKEL 205  
 QY 182 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 241  
 DB 206 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 265  
 QY 242 ATKAIGKRTGNNG-LEANSKNTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 299  
 DB 266 A-KAIGKRTGNNGSLDNEADHNGSLISGAVITSTLTQKLSKNGSEGLKEKIAAKKCESEFS 324  
 QY 300 TEFNTNKLSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTL 359  
 DB 325 TEFNTNKLSEHAVIGLIDNTDNOARAILKHKANKDKGAELFKFAVENLSKAODTL 384

OY 360 KNAVKELTS 368  
 DB 385 TNSVKELTS 393

## RESULT 14

US-09-596-746-32  
 : Sequence 32, Application US/09596746  
 : GENERAL INFORMATION:  
 : APPLICANT: Datwyler, Raymond J.  
 : APPLICANT: Selinost, Gerald  
 : APPLICANT: Dykhuizen, Dania  
 : APPLICANT: Luft, Benjamin J.  
 : APPLICANT: Maria J.C. Gomes-Solecki  
 : TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 : FILE REFERENCE: 2631.1002-001  
 : CURRENT APPLICATION NUMBER: US/09/596,746  
 : CURRENT FILING DATE: 2000-06-16  
 : PRIOR APPLICATION NUMBER: US 60/140,042  
 : PRIOR FILING DATE: 1999-06-18  
 : NUMBER OF SEQ ID NOS: 84  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 32  
 : LENGTH: 369  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: OspC Chimera  
 : US-09-596-746-32

Query Match 82.3%; Score 1479; DB 19; Length 369;  
 Best Local Similarity 83.8%; Pred. No. 4.9e-101;  
 Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

OY 2 ACNNSGKDGNASNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASIDELATKAI 61  
 DB 1 ACNNSGKDGNTSANSADSVKGPNTLEISKRTESNAVYLAVEREETLLASIDELATKAI 59  
 OY 62 GKRIKNGG-LEANOSKNTSLSGAVYISDLIAEKLNVKNEE-LKEKIDTAKOCSTEFN 119  
 DB 60 GKRIKNGGSLDNANRRESILAGAYITSTLTOKLSKLNCEGLKEKIAAKKCSSEFST 119  
 OY 120 KLSSEHAVGLDNLTDNNAORAILKKH-ANKDGAELERFRAVENLSKAADTLKNAV 178  
 DB 120 KLRDNHAGLGIGVTDENAKKAILKANAAGKDKGVELEKLSLSLSKAEMLANSV 179  
 OY 179 KELTSPVHGNNSRKDGNASTNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASI 238  
 DB 180 KELTSPVHGNNSRKDGNASTNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASI 239  
 OY 239 DELATRAIGKRIKNGGLEANSKNTSLSGAVYISDLIAEKLNVKNEELKEKIDTAKOC 298  
 DB 240 DELATRAIGKRIKNGGLEANSKNTSLSGAVYISDLIAEKLNVKNEELKEKIDTAKOC 299  
 OY 299 STEFTNKLKSEHAVGLDNLTDNNAORAILKKHANKDGAELERFRAVENLSKAADT 358  
 DB 300 STEFTNKLKSEHAVGLDNLTDNNAORAILKKHANKDGAELERFRAVENLSKAADT 359  
 OY 359 LKNAVKEKLS 368  
 DB 360 LKNAVKEKLS 369

## RESULT 15

US-09-596-746-68  
 : Sequence 68, Application US/09596746  
 : GENERAL INFORMATION:  
 : APPLICANT: Datwyler, Raymond J.  
 : APPLICANT: Selinost, Gerald  
 : APPLICANT: Dykhuizen, Dania

APPLICANT: Luft, Benjamin J.  
 : APPLICANT: Maria J.C. Gomes-Solecki  
 : TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
 : FILE REFERENCE: 2631.1002-001  
 : CURRENT APPLICATION NUMBER: US/09/596,746  
 : CURRENT FILING DATE: 2000-06-16  
 : PRIOR APPLICATION NUMBER: US 60/140,042  
 : PRIOR FILING DATE: 1999-06-18  
 : NUMBER OF SEQ ID NOS: 84  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 68  
 : LENGTH: 393  
 : TYPE: PRT  
 : ORGANISM: OspC Chimera  
 : US-09-596-746-68

Query Match 82.1%; Score 1476; DB 19; Length 393;  
 Best Local Similarity 83.5%; Pred. No. 8.9e-101;  
 Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

OY 2 ACNNSGKDGNASNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASIDELATKAI 61  
 DB 25 SCNNSGKDGNTSANSADSVKGPNTLEISKRTESNAVYLAVEREETLLASIDELATKAI 83  
 OY 62 GKRIKNGG-LEANOSKNTSLSGAVYISDLIAEKLNVKNEE-LKEKIDTAKOCSTEFN 119  
 DB 84 GKRIKNGGSLDNANRRESILAGAYITSTLTOKLSKLNCEGLKEKIAAKKCSSEFST 143  
 OY 120 KLSSEHAVGLDNLTDNNAORAILKKH-ANKDGAELERFRAVENLSKAADTLKNAV 178  
 DB 144 KLRDNHAGLGIGVTDENAKKAILKANAAGKDKGVELEKLSLSLSKAEMLANSV 203  
 OY 179 KELTSPVHGNNSRKDGNASTNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASI 238  
 DB 204 KELTSPVHGNNSRKDGNASTNSADESVKGPNTLEISKRTESNAVYLAVEREETLLASI 263  
 OY 239 DELATRAIGKRIKNGGLEANSKNTSLSGAVYISDLIAEKLNVKNEELKEKIDTAKOC 298  
 DB 264 DELATRAIGKRIKNGGLEANSKNTSLSGAVYISDLIAEKLNVKNEELKEKIDTAKOC 323  
 OY 299 STEFTNKLKSEHAVGLDNLTDNNAORAILKKHANKDGAELERFRAVENLSKAADT 358  
 DB 324 STEFTNKLKSEHAVGLDNLTDNNAORAILKKHANKDGAELERFRAVENLSKAADT 383  
 OY 359 LKNAVKEKLS 368  
 DB 384 LKNAVKEKLS 393

Search completed: March 18, 2002, 10:08:45  
 Job time: 973 sec

Tue Mar 19 10:57:53 2002

us-09-596-746a-38.rapm

Page 8

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw.model

Run on: March 18, 2002, 09:58:23 ; Search time 81.02 seconds

(Without alignments)  
290.158 Million cell updates/sec

Title: US-09-596-746a-38

Perfect score: 1797  
Sequence: 1 MACNNSGKDGNASANSADSE.....ENLSKRAODTLKNAVKELTS 368

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

## Database :

Pending Patents AA New: \*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	588.5	32.7	212	6	US-09-974-992-7 Sequence 7, Appl
2	583	32.4	211	6	US-09-974-992-5 Sequence 5, Appl
3	567	31.6	207	6	US-09-974-992-3 Sequence 3, Appl
4	567	31.6	207	6	US-09-974-992-3 Sequence 3, Appl
5	191	10.6	837	6	US-09-815-242-5883 Sequence 5883, Ap
6	191	10.6	837	7	US-10-072-851-5883 Sequence 5883, Ap
7	191	10.6	875	6	US-09-815-242-13080 Sequence 13080, A
8	191	10.6	875	7	US-10-072-851-13080 Sequence 13080, A
9	182	10.1	2434	6	US-09-815-242-5835 Sequence 5835, Ap
10	182	10.1	2434	7	US-10-072-851-5835 Sequence 5835, Ap
11	182	10.1	6281	6	US-09-815-242-12996 Sequence 12996, A
12	182	10.1	6281	7	US-10-072-851-12996 Sequence 12996, A
13	173.5	9.7	1881	7	US-10-032-585-7646 Sequence 7646, Ap
14	173.5	9.7	1881	7	US-10-032-585-7646 Sequence 7646, Ap
15	171.5	9.5	2076	6	US-09-815-242-5815 Sequence 5815, Ap
16	171.5	9.5	2076	6	US-09-815-242-5815 Sequence 5815, Ap
17	171.5	9.5	2186	6	US-10-072-851-12913 Sequence 12913, A
18	171.5	9.5	2186	6	US-10-072-851-12913 Sequence 12913, A
19	167	9.3	1144	6	US-09-708-427-15046 Sequence 15046, A
20	167	9.3	1144	6	US-09-708-427-15046 Sequence 15046, A
21	163.5	9.1	1304	6	US-09-708-427-15044 Sequence 15044, A
22	163.5	9.1	1304	6	US-09-708-427-15044 Sequence 15044, A
23	163.5	9.1	1686	6	US-09-708-427-13228 Sequence 13228, A
24	162.5	9.0	1730	6	US-09-708-427-13227 Sequence 13227, A
25	162.5	9.0	1014	6	US-09-708-427-19883 Sequence 19883, A
			1018	6	US-09-708-427-19882 Sequence 19882, A

26	162.5	9.0	1269	6	US-09-708-427-19881 Sequence 19881, A
27	162.5	9.0	2056	6	US-09-614-150-4824 Sequence 4824, Ap
28	159	8.8	2437	7	US-09-815-242-5834 Sequence 5834, Ap
29	159	8.8	2437	7	US-10-072-851-5834 Sequence 5834, Ap
30	158	8.8	5795	6	US-09-815-242-12610 Sequence 12610, A
31	158	8.8	5795	7	US-10-072-851-12610 Sequence 12610, A
32	157.5	8.8	2086	6	US-09-815-242-5639 Sequence 5639, Ap
33	157.5	8.8	2086	7	US-10-072-851-5639 Sequence 5639, Ap
34	157	8.7	1639	6	US-09-614-150-6213 Sequence 6213, Ap
35	156.5	8.7	1013	7	US-09-614-150-9139 Sequence 9139, Ap
36	152.5	8.5	1128	7	US-10-032-585-7627 Sequence 7627, Ap
37	152.5	8.5	1128	7	US-10-032-585-7627 Sequence 7627, Ap
38	152	8.5	2025	6	US-09-815-242-5703 Sequence 5703, Ap
39	152	8.5	2025	7	US-10-072-851-5703 Sequence 5703, Ap
40	152	8.5	3158	6	US-09-815-242-12611 Sequence 12611, A
41	152	8.5	3158	7	US-10-072-851-12611 Sequence 12611, A
42	151.5	8.4	2013	6	US-09-614-150-13758 Sequence 13758, A
43	149.5	8.3	1765	7	US-10-037-182-8 Sequence 8, Appl
44	149.5	8.3	1765	7	US-10-037-182-8 Sequence 8, Appl
45	149.5	8.3	1786	6	US-09-938-275-6 Sequence 6, Appl
			1786	6	US-09-873-676-113 Sequence 113, App

## ALIGNMENTS

RESULT 1  
US-09-974-992-7  
Sequence 7, Application US/09974/992  
GENERAL INFORMATION:  
APPLICANT: Mathiesen, Marianne J.  
APPLICANT: Thomsen, Michael  
APPLICANT: Holm, Arne  
APPLICANT: Ostergaard, Soren  
TITLE OF INVENTION: Novel OspC-derived peptide fragments  
FILE REFERENCE: 459-666p  
CURRENT APPLICATION NUMBER: US/09/974, 992  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: 09/180, 089  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: PCT/DK97/00203  
PRIOR FILING DATE: 1997-05-02  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 7  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia afzelii  
US-09-974-992-7

Query Match 32.7%; Score 588.5; DB 6; Length 212;  
Best Local Similarity 68.8%; Pred. No. 2.9e-32;  
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

1 MACNNSGKDGNASANSADSEYKGNLTREISKRTESNAVYIAVEVETLALASIDELATK 59  
17 ISCNNSGKGDSASNPDESAKGNLTREISKRTIDSNFAVAVETLVLSIDELATK 76  
60 AIGKRT-GNNGLEAENOSNTSLSCAVYISDLIAEKLVLNKL-EELKEKIDTAKCSTEE 117  
77 AIGKRTIDNNNGIALINNGSILACAVYISLTITREKLKLELKEIKAKKCSSEF 136  
118 TNKLSEHAVVLGDLNLTDDNORALITKHKANKCAAELEKLFKAVENLSKRAODTLNA 177  
137 TNKLKSHADGKQDADTDHAKAAILKTHATGDKAKKFEKDLFESEVGLKNAQVALTNS 196  
178 VKELTSPV 186  
197 VKELTSPV 205  
RESULT 2  
US-09-974-992-5

```
Sequence 5, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Thiesen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT FILING DATE: 2001-10-10
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 211
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-974-992-5
```

```
Query Match
Best Local Similarity 32.4%; Score 583; DB 6; Length 211;
Matches 129; Conservative 24; Mismatches 30; Indels 8; Gaps 5;

QY 1 MACNNSGKGNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKA 60
DB 17 ISCNNSGKGNNTSANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKA 75
QY 61 IGGKINGNGU---FANOSKNTSLSGAVAIISDLIAEKLNVKKEE-LKEKIDPAKOCSTE 116
DB 76 IGGKIKKNDGSGDANH--NBSLAGAVTISTLTOKLSTKNSSEKIKERIAAKKCSFE 133
QY 117 FTNKLKSEHAVLGLDNLTDNAGRAILKRH-ANKRKGAELEKLFKAVENLSRAADDTLK 175
DB 134 FSTKLKKNHNLQIGQVTDENAKKAILKANAAGKDKVELEKLSGSLSTRAAKEMLA 193
QY 176 NAYVELTSPV 186
DB 194 NSVKELTSPV 204
```

```
RESULT 3
US-09-974-992-3
Sequence 3, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Thiesen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT FILING DATE: 2001-10-10
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii
US-09-974-992-3
```

```
Query Match
Best Local Similarity 31.6%; Score 567; DB 6; Length 207;
Matches 118; Conservative 33; Mismatches 32; Indels 4; Gaps 3;
```

```
QY 1 MACNNSGKGNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKA 60
DB 17 ISCNNSG--GDSASTNDEBSAKGNPLTVISKITDSNAFLAVKEVALLSIDEL-SKA 73
QY 61 IGGKINGNG-LEANOSKNTSLSGAVAIISDLIAEKLNVKKEEIDPAKOCSTEFN 119
DB 74 IGGKIKKNDGTLDEANRANESLNGAATISKLITOKLSVLNSEEKKEIEAKDCSEFTT 133
QY 120 KLSSEHAVLGLDNLTDNAGRAILKRHANKRKGAELEKLFKAVENLSRAADDTLKNAV 179
DB 134 KLKDSHAELTIGQVODDNAAKAILKTHGTDKKGAKELEELFKSLSTSKAAQAALNSVK 193
QY 180 ELNTPV 186
DB 194 ELNTPV 200
```

```
RESULT 4
US-09-815-242-5883
Sequence 5883, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5883
LENGTH: 837
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5883
```

```
Query Match
Best Local Similarity 10.6%; Score 191; DB 6; Length 837;
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;
```

```
QY 4 NNSGKGNASANSADSVKGNLTKTESNAV--VLAVER-VELLASI----- 53
DB 100 NNNLEVAQNANFAIDLTSLNGPQRAKLEQYGQATTLNVCOTVRNQTTLWTAKGLR 159
QY 54 DELATKAIGKIKGNGNLEANSKNTSLSGAVAIISDLIAE-----KLVNKLNEELKEKIDT 109
DB 160 DSIANEAT-IRAGQNYDASQNKOTDYNASVTAARAKALIGQTTSPSMAQEIINAKDOV-T 217
QY 110 AKQCSSTFTNKLK-----SEHAVLGLDNLTDNAGRAILKRHANKRKGAELEKLFKAVE 164
DB 218 AKQQAALNGQENLTKTAQTNAAKHNLGSLDIT--AQKDAVKRQI---EGATHVNEVTOAON 272
```





DB 347 AYTNAVTOAQEOLINKAQCPNTSKDGEVETALENVQRAKNEELNGNONVANAKTTAKNALNL 406  
QY 276 IAEKLVNKEELKEKID-----TAKOCSTEFNKLSSEHAVLGLD-LTDD 321  
DB 407 TS--INNAQKALKSQIEGATTVAGVNOVSTTASELNTAMSN-----LONGINDE 454  
QY 322 NQORALTKHANKDGAEELEKFEKAVENLSKAAODTLTNAVEKELS 368  
DB 455 AATKALNCTONLEKAKQHANTAIIDGLSHLTNAQKALKOLVOQSTT 501

RESULT 7  
US-10-072-851-13080

; Sequence 13080, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits  
; TITLE OF INVENTION: Proliferation  
; FILE REFERENCE: ELITRA.028A  
; CURRENT APPLICATION NUMBER: US/10/072,851  
; CURRENT FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13080  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-072-851-13080

Query Match 10.6%; Score 191; DB 7; Length 875;  
Best Local Similarity 22.4%; Pred. No. 4e-05;  
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;

DB 4 NNSGKDNANSANSADE---SVKGNLTETISKITESNAV--VLAKE-VETLLASI----- 53  
QY 137 NNNLEVAKQNNANTAIIDGLTSLNGQRAKLEQVGQATTLTNPVQVRNMAOTLTNTAMGLR 156  
DB 54 DELATKAIGKKNNGLEANSQNTSLSGAYALSIDLIAE---KINVLKNEELKEKIDT 109  
QY 197 DSJANEAFT-IKAGQNTYDASQNKQTDYNSAVTAAKALIIGQTSPSMAQEIINAKQOV-T 254  
DB 110 AKOCSTEFNTKLR-----SEHAVLGLDNLTDNAQRAILKHKHANKDGAEELEKFEAVE 164  
QY 255 AKQOALNGQENLRTAQTNAKQHLNGLSLTD--AQKDAVKRQI--EGATHVNEVTOAQN 309  
DB 165 NLS--KAAQDTLKAAYKELTSPYHGNNSRKDNASTNSADESVKGNLTETISKITESN 222  
QY 310 NADALNTAMTNRKNGID-----QNTIKQGS-VNFTDDAE-----KRN 346  
DB 223 AVVLAKEVEETLL-----ASIDELATKAIGKKNNGLEANSQNTSLSGAYALSIDL 275  
QY 347 AYTNAVTOAQEOLINKAQCPNTSKDGEVETALENVQRAKNEELNGNONVANAKTTAKNALNL 406  
DB 276 IAEKLVNKEELKEKID-----TAKOCSTEFNKLSSEHAVLGLD-LTDD 321  
DB 407 TS--INNAQKALKSQIEGATTVAGVNOVSTTASELNTAMSN-----LONGINDE 454

QY 322 NQORALTKHANKDGAEELEKFEKAVENLSKAAODTLTNAVEKELS 368  
DB 455 AATKALNCTONLEKAKQHANTAIIDGLSHLTNAQKALKOLVOQSTT 501

RESULT 8  
US-09-815-242-5835

; Sequence 5835, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5835  
; LENGTH: 2434  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5835

Query Match 10.1%; Score 182; DB 6; Length 2434;  
Best Local Similarity 22.2%; Pred. No. 0.00057;  
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

QY 4 NNSGKDNANSANSADE---SVKGNLTETISKITESNAV--VLAKE-VETLLASI----- 53  
DB 702 NNNLEVAKQNNANTAIIDGLTSLNGQRAKLEQVGQATTLTNPVQVRNMAOTLTNTAMGLR 761  
QY 54 DELATKAIGKKNNGLEANSQNTSLSGAYALSIDLIAE---KINVLKNEELKEKIDT 109  
DB 762 DSJANEAFT-IKAGQNTYDASQNKQTDYNSAVTAAKALIIGQTSPSMAQEIINAKQOV-T 819  
QY 110 AKOCSTEFNTKLR-----SEHAVLGLDNLTDNAQRAILKHKHANKDGAEELEKFEAVE 157  
DB 820 AKQOALNGQENLRTAQTNAKQHLNGLSLTD--AQKDAVKRQI--EGATHVNEVTOAQNND 877  
QY 158 KLEKAVENLSKAAQD--TLK-----NAVKELTSPY--VHGNNSRKD-- 194  
DB 878 ALNTAMTNRKNGIDQNTIKQGVNFTDADAEKRNATYNTAVTOAQEOLINKRAQCPNTSKDGV 937  
QY 195 -----GNASTNSADESVKGP--NLTEISKITES-----NA 223  
DB 938 ETALENVQRAKNEELNGNONVANAKTTAKNALNLINNAQKALKSQIEGATTVAGVNO 997  
QY 224 VVLAKEVEETLLASI-----DELATKAIGKKNNGLEANSQNTSLSGAYALSIDLI-- 276  
DB 998 VSTTASELNTAMSNLONGINDEAATKAQKY-----TDADREKQTVYNDNAVTAKTLLDK 1052

OY 277 -----AEKLVNFK-----NEELKEKIDTAQO-----CSTEFNTKL 306  
DB 1053 TAGSNDKNKAVEQALORVNTAKTALNGDERLNEAKNTAKQOAVTMSHLTDAOKANLTSQI 1112  
OY 307 KSEHAVIGLIDNT-----TDDNAORAILKKNANKDKGAELKLEKFAVENLSKAQODTLKN 361  
DB 1113 ESGTIVAGVQIOANGTLDOAMNOLROSIASND-ATKSSSEYODANADLQNAVDNAYTN 1171  
OY 362 A 362  
DB 1172 A 1172

RESULT 9  
US-10-072-851-5835  
; Sequence 5835, Application US/10072851  
; GENERAL INFORMATION:

APPLICANT: Carr, Grant J.  
APPLICANT: Xu, H. Howard  
APPLICANT: Foulkes, J. Gordon  
APPLICANT: Zamudio, Carlos  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Roemer, Terry  
APPLICANT: Boone, Charles  
APPLICANT: Bussey, Howard  
TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits  
FILE REFERENCE: ELITRA.028A  
CURRENT APPLICATION NUMBER: US/10/072,851  
CURRENT FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 15811  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO: 5835  
LENGTH: 2434  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-10-072-851-5835

Query Match 10.1%; Score 182; DB 7; Length 2434;  
Best Local Similarity 22.2%; Pred. No. 0.00057;  
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

OY 4 NNSGKDGNANSADSE---SVKGNPLTEISKITESNAV--VLAWE-VETLLASI----- 53  
DB 702 NNNLEVAQONANTRAIIDGLTSLNGPQAKLKEQVQATTLTPNVQTVDRDNGQTLNTAKGLR 761  
OY 54 DELATKAIGKKNNGLEANSQKNTSLSGAYAISDLIAE---KLVNKLNEELKEKIDT 109  
DB 762 DSIANEAT-IKAGQNTYDASQNKOTDYNSAVTAKAIIIGTSPSMAQOEINQAKDOV-T 819  
OY 110 AKOCSTEFNTKKL-----SEHAVIGLIDNTDDNAORAILK-----HAK-OKGAEELE 157  
DB 820 AKQOALNGOENLRTAQTNAKOHNLGLSLDLD--AQKDAVKRQEGATHVEVYVQAQNNAD 877  
OY 158 KLEKAVENLSKAQD--TLK-----NAVVELTSPI--VHGNSNRKD-- 194  
DB 878 ALNTAMTINLKNIGIDQNTIKQGVNFTDADBAKRNAYTNNAVTOAEQILNKRAQGNPSKDV 937  
OY 195 -----GNASTNSADESVKCP--NLTEISKITES-----NA 223  
DB 938 ETALENVORAKNELNGNQNANAKKTAKNALNLTSINNAQKALKSQIEGATTVAGVNO 997  
OY 224 VVLAWEVETLLASI-----DELATKAIGKKNNGLEANSQKNTSLSGAYAISDLI-- 276

DB 998 VSTTASBELNTAMNSLNGINDENATKAQKY-----TDADREKQTAAYNDVATKATLDD 1052  
OY 277 -----AEKLVNFK-----NEELKEKIDTAQO-----CSTEFNTKL 306  
DB 1053 TAGSNDKNKAVEQALORVNTAKTALNGDERLNEAKNTAKQOAVTMSHLTDAOKANLTSQI 1112  
OY 307 KSEHAVIGLIDNT-----TDDNAORAILKKNANKDKGAELKLEKFAVENLSKAQODTLKN 361  
DB 1113 ESGTIVAGVQIOANGTLDOAMNOLROSIASND-ATKSSSEYODANADLQNAVDNAYTN 1171  
OY 362 A 362  
DB 1172 A 1172

RESULT 10  
US-09-815-242-12996  
; Sequence 12996, Application US/09815242  
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO: 12996  
LENGTH: 6281  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12996

Query Match 10.1%; Score 182; DB 6; Length 6281;  
Best Local Similarity 22.2%; Pred. No. 0.0019;  
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

OY 4 NNSGKDGNANSADSE---SVKGNPLTEISKITESNAV--VLAWE-VETLLASI----- 53  
DB 1481 NNNLEVAQONANTRAIIDGLTSLNGPQAKLKEQVQATTLTPNVQTVDRDNGQTLNTAKGLR 1540  
OY 54 DELATKAIGKKNNGLEANSQKNTSLSGAYAISDLIAE---KLVNKLNEELKEKIDT 109  
DB 1541 DSIANEAT-IKAGQNTYDASQNKOTDYNSAVTAKAIIIGTSPSMAQOEINQAKDOV-T 1598  
OY 110 AKOCSTEFNTKKL-----SEHAVIGLIDNTDDNAORAILK-----HAK-DKGAEELE 157  
DB 1599 AKQOALNGOENLRTAQTNAKOHNLGLSLDLD--AQKDAVKRQEGATHVEVYVQAQNNAD 1656  
OY 158 KLEKAVENLSKAQD--TLK-----NAVVELTSPI--VHGNSNRKD-- 194

```

Db 1657 ALNTAMTNLKNIGDONTIKOGVNFDTDADEAKRNAYTNVTAQAEQILNKAAQPTSKGV 1716
OY 195 -----GNASTNSADESVKGP--NLTEISKITTES-----NA 223
Db 1717 ETALENVORAKNELNGONAVNAKTTAKNALNNLTSSINNAOKELAKSQIEGATTVAGVNO 1776
OY 224 VILAAREVEETLLASI-----DELATKAIGKIGNGNGLEANOSKNTSLSGAVALISDLI-- 276
Db 1777 VSTTASELNTAMSNLQNGINDEAATKAQKY-----TDAREKOTAYNDAYTAKTLLDK 1831
OY 277 -----AEKLVK-----NEELKEKIDTAQO-----CSTEFNKL 306
Db 1832 TAGSNDNKAAREQALORVNTAKTALNGDERLNEAKNTAKQOVATWSHLTDAQKANLTSGI 1891
OY 307 KSEHAVLGLDNL-----TDDNAGRAILKHNKDKGAELKFKRAVENLSKAQDTLKN 361
Db 1892 ESGTTVAGVGIQANAGTLDQANMOLROSIAKSD--ATKSSDYODANADLQANAYNDAYTN 1950
OY 362 A 362
1951 A 1951

```

RESULT 11  
US-10-072-851-12996

Sequence 12996, Application US/10072851

```

GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C
TITLE OF INVENTION: Proliferation
FILE REFERENCE: ELITRA 028A
CURRENT APPLICATION NUMBER: US/10/072.851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FASTSEQ for Windows Version 4.0
NO ID NO 12996
LENGTH: 6281
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-12996

```

Query Match Best local Similarity 10.1%; Score 182; DB 7; Length 6281;

Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

```

OY 4 NNSGKDNASANSAD--SVKGNLTETISKITESNAV--VLAVKE-VEETLLASI----- 53
Db 1481 NNNLEVAKNONANTAIIDGITSINGQOKAKLEQVGOATTLTPVOTVRONAOITLNTAMKGLR 1540
OY 54 DELATKAIGKIGNGNGLEANOSKNTSLSGAVALISDLIAE-----KLNVLKNEELKEKIDT 109
Db 1541 DSIINEAT-IKAGONYDASQNKQTDYNSAVTAKAKIIGQTFSPSMAQELNQAQDY-T 1598
OY 110 AKOOSTEFTNKL-----SEHAVLGLDNLDDNAGRAILK-----HANK-DKGAEELE 157
Db 1599 AKQALNGOENLRTAQTNAAQHLNGLSLDITD--AQKDAVAVROIEGATVHNEVTAQANMAD 1656

```

```

OY 158 KLFKAVENLSKAQD--TLK-----NAVKEITSPH--VHGNSNRKD-- 194
Db 1657 ALNTAMTNLKNIGDONTIKOGVNFDTDADEAKRNAYTNVTAQAEQILNKAAQPTSKGV 1716
OY 195 -----GNASTNSADESVKGP--NLTEISKITTES-----NA 223
Db 1717 ETALENVORAKNELNGONAVNAKTTAKNALNNLTSSINNAOKELAKSQIEGATTVAGVNO 1776
OY 224 VILAAREVEETLLASI-----DELATKAIGKIGNGNGLEANOSKNTSLSGAVALISDLI-- 276
Db 1777 VSTTASELNTAMSNLQNGINDEAATKAQKY-----TDAREKOTAYNDAYTAKTLLDK 1831
OY 277 -----AEKLVK-----NEELKEKIDTAQO-----CSTEFNKL 306
Db 1832 TAGSNDNKAAREQALORVNTAKTALNGDERLNEAKNTAKQOVATWSHLTDAQKANLTSGI 1891
OY 307 KSEHAVLGLDNL-----TDDNAGRAILKHNKDKGAELKFKRAVENLSKAQDTLKN 361
Db 1892 ESGTTVAGVGIQANAGTLDQANMOLROSIAKSD--ATKSSDYODANADLQANAYNDAYTN 1950
OY 362 A 362
1951 A 1951

```

RESULT 12  
US-10-032-585-7646

Sequence 7646, Application US/10032585

```

GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032.585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7646
LENGTH: 1881
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1881)..(1881)
OTHER INFORMATION: x-any amino acid
US-10-032-585-7646

```

Query Match Best local Similarity 9.7%; Score 173.5; DB 7; Length 1881;

Matches 112; Conservative 55; Mismatches 150; Indels 121; Gaps 25;

```

OY 2 ACNNSGRD-----GNASANSAD--ESVK-----GPNLTETISKITESNAVVLAVKE 45
Db 1470 ALTRKSSKDEVCNOKSELODSLVKSVSEKLNENKYNQETSLKMEIEKO-----KE 1523
OY 46 VETLLASIDELAKKAIGKIGNGNGLEANOSKNTSLSGAVALISDLI--AEKLVKNEE 102
Db 1524 IYTLQETELKRISEVERK-----AMLSENSEVIKEY--SDKIKSLSEKINSIKENH 1574
OY 103 LKEKIDTAQOSTEFTNKLKSEHAVLGLDN-----LTDNAG--RAILKHH----- 146
Db 1575 SKK-ITTHNEQKT-----SLKQDIKLSQDHESQVQLEDEKQNLKELKASLEKHNHTESAT 1629
OY 147 --ANKDKGAELKFKRAVENLSKAQDTLKNAYVELTSPYVHGNSNRKDNASTNS-- 201
Db 1630 SLEEKNOJIKELSETIKSLKTELKTSGDALKOSQKEY-----KLTKTKNSDRESKLE 1681
OY 202 -----ADESVKGNLTETISKITESNAVVLAVKEVEETLLASI-----LASTIDELA 242
Db 1682 KOLELEKYSLOLQINDEKLTGITERETALK-----SELETYKNSGLSTTSELA 1730

```

243 TRAIKRIKNGNGLEANOSKNTSLSG--AYAIIDL-----IAKLVNKLNELEK--- 291  
1731 --ALRTV--KSLE-KEKEELOFLSGNKSKELEDYIOKHSIDSEKIKAL--TDELKEKTKQ 1784  
292 IDTAOCSTEFNTKLKSEHAVALGLDNLTDNORAILK--HANKDKGALEKLFKAVERN 350  
1785 FDSCKKLTLENDLSTK-----KELETKQTSTKFKNLEERKDEIYKLVNKELELLKN 1839  
351 LSKAODTLKNAVKELTS 368  
1840 DMSGAKKELSEKVS KLES 1857

RESULT 13  
US-10-072-851-15590

Sequence 15590, Application US/10072851  
GENERAL INFORMATION:

APPLICANT: Carr, Grant J.  
APPLICANT: Xu, H. Howard  
APPLICANT: Foulkes, J. Gordon  
APPLICANT: Zamudio, Carlos  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Roemer, Terry  
APPLICANT: Jiang, Bo  
APPLICANT: Boone, Charles  
APPLICANT: Bussey, Howard  
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C  
FILE REFERENCE: ELITRA.028A  
CURRENT APPLICATION NUMBER: US/10/072,851  
CURRENT FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 1581  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15590  
LENGTH: 1881  
TYPE: PRN  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1881) (1881)  
OTHER INFORMATION: X-any amino acid  
US-10-072-851-15590

Query Match 9.7%; Score 173.5; DB 7; Length 1881;  
Best Local Similarity 25.6%; Pred. No. 0.0015;  
Matches 112; Conservative 55; Mismatches 150; Indels 121; Gaps 25;

2 ACNNSGKD-----GNASANSAD--ESVK-----GPNLTESKITESNAVLAVER 45  
1470 ALTSSKDLLEVGNQSGNKSSELSKFNKFNQDETSLADEIKQ-----KE 1523  
46 VETLASIDELATKAIGKRIKNGNGLEANOSKNTSLSGAYAIIDL---AEKLVNKLNEE 102  
1524 IYVLTOTELKDRISEVEKER-----AMSENSEYIKEY--SDRIKSLESKINSIKENH 1574  
103 LKEKIDTAOCSTEFNTKLKSEHAVALGLDNLTDNNO-----RAILKHH----- 146  
1575 SKF-ITTTNEOKT-----SLKODIAKLSODHESAQOTULEDKENQLEKLSLEKHNTEGAT 1629  
147 --ANKDKGALEKLFKAVERNLSKAADTLKNAVKELTSPIVHGNNSRKDGNASTNS--- 201  
1630 SIEEKNNQIKELSETIKSLKTELKTSGDALQSQKEY-----KTLTKNSDSTESKLE 1681  
202 -----ADESVAGPNLTETISKITESNAVLAVERETL-----LASIDELA 242

1682 KOLELEKYSQDLQTADEKLTGTERETALK-----SELETVYVNSGLSTSELA 1730  
243 TRAIKRIKNGNGLEANOSKNTSLSG--AYAIIDL-----IAKLVNKLNELEK--- 291  
1731 --ALRTV--KSLE-KEKEELOFLSGNKSKELEDYIOKHSIDSEKIKAL--TDELKEKTKQ 1784  
292 IDTAOCSTEFNTKLKSEHAVALGLDNLTDNORAILK--HANKDKGALEKLFKAVERN 350  
1785 FDSCKKLTLENDLSTK-----KELETKQTSTKFKNLEERKDEIYKLVNKELELLKN 1839  
351 LSKAODTLKNAVKELTS 368  
1840 DMSGAKKELSEKVS KLES 1857

RESULT 14  
US-09-815-242-5815

Sequence 5815, Application US/09815242  
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5815  
LENGTH: 2076  
TYPE: PRN  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5815

Query Match 9.5%; Score 171.5; DB 6; Length 2076;  
Best Local Similarity 20.0%; Pred. No. 0.0023;  
Matches 85; Conservative 78; Mismatches 174; Indels 87; Gaps 14;

20 SVKGNLTETISKITESNAVVL--AVKEVETLLASIDELATKA---IGKRIKNGGLEA-- 72  
1132 TAKKOEINONTATTEDEKOVALLNQVDEELATAINNINQADTNAEVDQAOLGTKAINAIO 1191  
73 -----NOSKNISL-----SGAYAIIDLKLVNKLNELEKIDTAOCSTEF 116  
1192 PNIVKPAALAOINOHYNAKLAELINTPDTYNDKNAAINTL--NODROQATIESIKQAN-- 1248  
117 ETNKLSEHAVALGLDNLTDNORAILK--HANKDKGALEKLFKAVERNLSKAADTLKN 176  
1249 --TNAEVDQAAYTAENNI--DAVGVVYVKKQARDKITAFAKRIEAVKQPNATDEKQNA 1305  
177 AVKELTS-----PIVHGNNSRKDGNASTN-----SADESVAGPNLTETISKITES 221

```

Db 1306 AVNQINOLKDAQAINQONQNDQVDTTNOAVNAIDNVEAEVYIKPRAIDIEKAVKX 1365
OY 222 N-----AVVLAVKEVEETLLASIDELATKAIGKIGNNGLEANSKNTSL 265
Db 1366 QOQIDNSIDSTDNEKEVASQALAKEKEKALAIIDQAOQNSQVNOAATNGVSA-----IKI 1420
OY 266 LSGAVAISDLIAEKLNVKNEELKEKIDTAOKCS-----TEPTNKIKSEHAVIG 314
Db 1421 IQPETHKVPAREKIN-QKANELRAKINODKEATAEERQVALDKINEFVNQAMTDITNKR 1479
OY 315 LDNLTDDNAQRAI-----LKKHANKDKGAAELEKLEKRAVENLSKAADTLKNAVK 364
Db 1480 TNOQVDDTTSQALDSIALVTPDHIIVRAAADVAQOYEAKKREIEQAEHATDEEKQVALN 1539
OY 365 ELTS 368
Db 1540 QLAN 1543

```

```

BLT 15
Sequence 5815, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C
TITLE OF INVENTION: Proliferation
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ. ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5815
LENGTH: 2076
TYPE: PRT
ORGANISM: Staphylococcus aureus
DB 10-072-851-5815

```

Query Match 9.5%; Score 171.5; DB 7; Length 2076;  
 Best Local Similarity 20.0%; Pred. No. 0.0023;  
 Matches 85; Conservative 78; Mismatches 174; Indels 87; Gaps 14;

```

OY 20 SVKGPNLTEISKITTESNAVVL-AVKEVEETLLASIDELATKA--IGKIGNNGLEA-- 72
Db 1132 TAKQOEINQNTNATTEERQVALNOVDOLATININNQADTNAEYDQAOQIGTKAINAIQ 1191
OY 73 -----NOSKNTSL--SGAYAISDLIAEKLNVKNEELKEKIDTAOKCSTE 116
Db 1192 PNIVKKPAALAIQINQHNNAKLAEINATPDATNDEKNAINTL-NODRQOAIESTIKQAN-- 1248
OY 117 FTNKLKSEHAVIGLNDLTDDNAQRAILKHKANKDKGAAELEKLEKRAVENLSKAADTLK 176
Db 1249 -TNAEVDQAATVAENNI--DAVQYDVYVKKQAAARDKITAEVAKRIEAVQKQTPNATDEEKQ 1305
OY 177 AVKEITS-----PIVGNNSRKDGMASTN-----SADESYYKGNLTETISKITES 221
Db 1306 AVNQINOLKDAQAINQONQNDQVDTTNOAVNAIDNVEAEVYIKPRAIDIEKAVKX 1365

```

```

OY 222 N-----AVVLAVKEVEETLLASIDELATKAIGKIGNNGLEANSKNTSL 265
Db 1366 QOQIDNSIDSTDNEKEVASQALAKEKEKALAIIDQAOQNSQVNOAATNGVSA-----IKI 1420
OY 266 LSGAVAISDLIAEKLNVKNEELKEKIDTAOKCS-----TEPTNKIKSEHAVIG 314
Db 1421 IQPETHKVPAREKIN-QKANELRAKINODKEATAEERQVALDKINEFVNQAMTDITNKR 1479
OY 315 LDNLTDDNAQRAI-----LKKHANKDKGAAELEKLEKRAVENLSKAADTLKNAVK 364
Db 1480 TNOQVDDTTSQALDSIALVTPDHIIVRAAADVAQOYEAKKREIEQAEHATDEEKQVALN 1539
OY 365 ELTS 368
Db 1540 QLAN 1543

```

Search completed: March 18, 2002, 09:58:28  
 Job time: 386 sec

Tue Mar 19 10:57:53 2002

us-09-596-746a-38.rapn

Page 9

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:48 ; Search time 68.77 Seconds  
(without alignments)  
407.623 Million cell updates/sec

Title: US-09-596-746a-38

Perfect score: 1797  
Sequence: 1 MACNNSGKDGMSANSADSES.....ENLSKAQDTLKNVKEITS 368

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.68:\*

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	50.1	209	2	I40142
2	661	36.8	193	2	S70287
3	646	35.9	211	2	I40145
4	642	35.7	193	2	S70279
5	639	35.6	191	2	I40153
6	623.5	34.7	210	2	G70218
7	622	34.6	209	2	S69917
8	619.5	34.5	210	2	S69927
9	619	34.4	193	2	S70286
10	614	34.2	191	2	S70284
11	612.5	34.1	212	2	S69922
12	605.5	33.7	194	2	S70268
13	605	33.7	178	2	I40104
14	603.5	33.7	178	2	I40125
15	603.5	33.6	210	2	I40144
16	596.5	33.2	212	2	I40143
17	589	32.8	193	2	S70274
18	588.5	32.7	212	2	S20543
19	588	32.7	211	2	I40268
20	587	32.7	211	2	S69918
21	586	32.6	211	2	S69932
22	584.5	32.5	212	2	S70254
23	581	32.3	193	2	S70280
24	579.5	32.2	212	2	S69921
25	577	32.1	211	2	I40278
26	575	32.0	193	2	S70265
27	572	31.8	193	2	S70276
28	569.5	31.7	177	2	I40129
29	569	31.7	207	2	I40271

30	567	31.6	200	2	I40122	outer surface prot
31	567	31.6	207	2	S69919	outer surface prot
32	567	31.6	207	2	S77727	outer surface prot
33	566.5	31.5	190	2	S70261	outer surface prot
34	566.5	31.5	194	2	S70277	outer surface prot
35	566.5	31.5	212	2	I40279	outer surface prot
36	566	31.5	209	2	I40281	outer surface prot
37	565.5	31.5	194	2	S70289	outer surface prot
38	565	31.4	211	2	I40277	outer surface prot
39	564	31.4	207	2	I40276	outer surface prot
40	562.5	31.3	214	2	S69916	outer surface prot
41	562	31.3	209	2	I40285	outer surface prot
42	560.5	31.2	180	2	I40110	outer surface prot
43	559	31.1	189	2	S70263	outer surface prot
44	559	31.1	203	2	I40108	outer surface prot
45	558	31.1	178	2	I40124	outer surface prot

## ALIGNMENTS

RESULT 1  
I40142  
outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40142; S70283  
R:Stevenson, B.; Bockenstedt, L.K.; Barthold, S.W.  
Infect. Immun. 62, 3568-3571, 1994  
A:Title: Expression and gene sequence of outer surface protein C of Borrelia burgdorferi  
A:Reference number: I40142; MUID:94314484  
A:Accession: I40142  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-209 <RES>  
A:Cross-references: EMBL:U04240; NID:92314877; PIDN:AACA5538.1; PID:9434658  
R:Liway, I.; Gibbs, C.P.; Schuster, R.; Dornet, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease spirochetes  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70283  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-209 <LIV>  
A:Cross-references: EMBL:L42894; NID:9858722; PIDN:AAB37002.1; PID:91695219  
A:Experimental source: strain 28691  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.1%; Score 901; DB 2; Length 209;  
Best Local Similarity 98.9%; Pred. No. 2, 2e-35;  
Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1	MACNNSGKDGMSANSADSESVKPNITEISKTTESNAVYLVAVKEVETLLASIDELATKA 60
DB	17	ISCNNSGKDGMSANSADSESVKPNITEISKTTESNAVYLVAVKEVETLLASIDELATKA 76
QY	61	IGKKITNNGLNENQSKNTSLSGAVVISPDIKLVNKKKEELKEKIDTAQKOSTEPTNK 120
DB	77	IGKKITNNGLNENQSKNTSLSGAVVISPDIKLVNKKKEELKEKIDTAQKOSTEPTNK 136
QY	121	LKSEHAVLGLDNLTDNAQRAILKKRANKDGAEELEKLFKAVENLSKAQDTLKNVKE 180
DB	137	LKSEHAVLGLDNLTDNAQRAILKKRANKDGAEELEKLFKAVENLSKAQDTLKNVKE 196
QY	181	LTSPIV 186
DB	197	LTSPIV 202
RESULT	2	

S70287

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C:Accession: S70287

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70287

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-193 &lt;LIV&gt;

A:Cross-references: EMBL:U042895; NID:9858723; PIDN:AA37003.1; PID:g1695220

A:Experimental source: strain 28354

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.8%; Score 661; DB 2; Length 193;

Best Local Similarity 74.7%; Pred. No. 2.3e-24;

Matches 139; Conservative 20; Mismatches 25; Indels 2; Gaps 2;

QY 3 CNGSGKDNASANSADSVKGNLTETSKRTESNAVLAKEVEYTLASIDELATKAI 62

Db 1 CNGSGKDNASANSADSVKGNLTETSKRTESNAVLAKEVEYTLASIDELATKAI 60

QY 63 KRTI-GNNGLEANOSKNTSLSGAVASDLIAEKLNVKLN-EELKEKIDPAKOCSTEF 120

Db 61 KRTI-GNNGLEANOSKNTSLSGAVASDLIAEKLNVKLN-EELKEKIDPAKOCSTEF 120

QY 121 LKSEHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAVE 180

Db 121 LKSEHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAVE 180

QY 121 LKSEHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAVE 180

Db 121 LKSEHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAVE 180

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 181 LTSPV 186

Db 181 LTSPV 186

QY 119 NKLSHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAV 178

Db 137 NKLSHAVLGIDNLTDDNAQRALILKHKNDKGAELKFKAVENSKAADTLKNAV 196

QY 179 KELTSPIV 186

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

Matches 136: Conservative 20: Mismatches 26: Indels 2: Gaps 2:

QY 5 NSGKDNASANSADSVKGNLPEISKRITESNAVLAKEVEETLLASIDELATKAIGK 64  
1 NSGKGNSTANSADSVKGNLPEISKRITESNAVLAKEVEETLLASIDELATKAIGK 60

QY 65 I-GNNGLEANSKNTSLISGAYAISDLIAEKLAVLNK-ELEKEIDTAQOCSTEFNTKK 122  
61 IONGSLAEAGHNGTLLAGAVYISKLITQKLDGKNESEKLEKEIENAKKCSDEFTKLE 120

QY 123 SEAVAVGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAODTLKNAVKELT 182  
121 GEMAOGLIEVNTDENAKKAILITDAKDKGAELKLFKAVENLSKAOKEMLANSVKELT 180

QY 183 SPVIV 186  
181 SPVIV 184

RESULT 6  
670218  
Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 13-Feb-1998 #sequence, revision 13-Feb-1998 #text\_change 21-Jul-2000  
C/Accession: G70218, I40269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Kierlavage, A.R.; Queckenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: G70218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BBB19  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A>Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la  
A:Reference number: I40269; MUID:96025162  
A:Accession: I40269  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49497; NID:g707092; PIDN:BA08457.1; PID:g769684  
R:Jaurs-Helke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A>Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)  
A:Reference number: S37726; MUID:93268136  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A>Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea  
A:Reference number: S70255; MUID:96266448  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:I42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212  
A:Experimental source: strain ip2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 34.7%; Score 623.5; DB 2; Length 210;  
Best Local Similarity 69.5%; Pred. No. 1.3e-22;

Matches 130: Conservative 24: Mismatches 32: Indels 1: Gaps 1:

QY 1 MACNNSGKDNASANSADSVKGNLPEISKRITESNAVLAKEVEETLLASIDELATKA 60  
17 ISCNNSGKDNSTANSADSVKGNLPEISKRITDSNVAVLVAKEVEETLLASIDELATKA 76

QY 61 IGKRI-GNNGLEANSKNTSLISGAYAISDLIAEKLAVLNK-ELEKEIDTAQOCSTEFNT 119  
77 IGKRIHONGNIDTENNNGSLLAGAVYSTLIKQKLDGKNESEKLEKEIDAKKCSDEFTTN 136

QY 120 KLRSEHVLGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAODTLKNAVK 179  
137 KLRKHVDLGEVYTDADAKAAILKTNGTKTGABELKLFESYEVLSKAOKEMLANSVK 196

QY 180 LTSPVIV 186  
197 LTSPVIV 203

RESULT 7  
569917  
outer surface protein C precursor - Lyme disease spirochete (strain B. pacificus)  
C/Species: Borrelia burgdorferi (Lyme disease spirochete)  
A:Variety: strain B. pacificus  
C>Date: 06-Dec-1996 #sequence, revision 14-Feb-1997 #text\_change 26-May-2000  
C/Accession: S69917; S72679  
R:Jaurs-Helke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A>Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69917  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-209 <JAU>  
A:Cross-references: EMBL:X83555  
A:Experimental source: strain B. pacificus  
R:Roessler, D.  
submitted to the EMBL Data Library, January 1995  
A:Reference number: S72679  
A:Accession: S72679  
A:Molecule type: DNA  
A:Residues: 1-50, 'E', '52-61', 'I', '63-117', 'D', '119-122', 'V', '124-173', 'D', '175-176', 'D', '178-209  
A:Cross-references: EMBL:X83555; NID:g872019; PIDN:CA58545.1; PID:g872020  
C:Superfamily: Lyme disease spirochete surface protein C  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-209/Product: outer surface protein C #status predicted <MAT>

Query Match 34.6%; Score 622; DB 2; Length 209;  
Best Local Similarity 69.9%; Pred. No. 1.6e-22;  
Matches 130: Conservative 24: Mismatches 32: Indels 0: Gaps 0:

QY 1 MACNNSGKDNASANSADSVKGNLPEISKRITESNAVLAKEVEETLLASIDELATKA 60  
17 ISCNNSGKDNSTANSADSVKGNLPEISKRITDSNVAVLVAKEVEETLLASIDELATKA 76

QY 61 IGKRI-GNNGLEANSKNTSLISGAYAISDLIAEKLAVLNK-ELEKEIDTAQOCSTEFNT 120  
77 IGKRIIDANGVQVQNGSLLAGAVYSTLIKQKLDGKNESEKLEKEIDAKKCSDEFTTN 136

QY 121 KLRSEHVLGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAODTLKNAVK 180  
137 LKNGNAOGLAAATDDNNAKAILKTNGTNDGAKELKELSESYSVLYAAQVMTLNSVKE 196

QY 181 LTSPVIV 186  
197 LTSPVIV 202

RESULT 8  
569927  
outer surface protein C precursor - Lyme disease spirochete (strain Pk6)  
C/Species: Borrelia burgdorferi (Lyme disease spirochete)

A:Variety: strain PKa  
C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69927; S72669  
R:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, H.  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia*  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69927  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69589  
R:Jauris, S.  
A:Experimental source: strain PKa  
submitted to the EMBL Data Library, February 1994  
A:Reference number: S72669  
A:Accession: S72669  
A:Molecule type: DNA  
A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAU>  
A:Cross-references: EMBL:X69589  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 34.5%; Score 619.5; DB 2; Length 210;  
Best Local Similarity 69.0%; Pred. No. 2.1e-22;  
Matches 129; Conservative 25; Mismatches 32; Indels 1; Gaps 1;

QY 1 MACNSGKDNANASANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKA 60  
DB 17 ISCNNSGKDNNTSANSANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKA 76  
QY 61 IGKRI-GNNLEANSQNTSLSGAVASDLIAEKLVNLKNEELKRIIDTAKOCSTFTN 119  
DB 77 IGKRIHNNGLDENNNHSLAGAVASFLIKIDGLKNEGLKEKIEAKKCSFTFTN 136  
QY 120 KLSKHAVALGLDNLTDNAORAILKKHANKDGAELKLFKAVENTLSKRAADTLKNAVK 179  
DB 137 KLMDKHTDLGKEGVTDADAKAILKTKNGTKTGAEELGKLFESVLSKRAKEMLSNVK 196  
QY 180 ELTSPV 186  
DB 197 ELTSPV 203

QY 9  
DB 86

Outer surface protein C - Lyme disease spirochete  
C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70286  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in *OspC* variation in Lyme disease  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70286  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:LA2897; NID:9858728; PIDN:AAB37006.1; PID:91695222  
A:Experimental source: strain 26815  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 34.4%; Score 619; DB 2; Length 193;  
Best Local Similarity 71.1%; Pred. No. 2e-22;  
Matches 133; Conservative 22; Mismatches 28; Indels 4; Gaps 3;

QY 3 CANSKGDNANASANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKAIG 62  
DB 11 CANSKGDNANASANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKAIG 62

DB 1 CANSKGDNANSTNSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKAIG 59  
QY 63 KKGNNGLNANQSKNTSLSGAVASDLIAEKLVNLKNEELKRIIDTAKOCSTFTNKL 121  
DB 60 KGINNGLDVOVNNASLAGAHTISKLVTESKILKNSGLKEKIEDKACSDFTTKL 119  
QY 122 KSEHAVLGLDNLTDNAORAILKKHANKDGAELKLFKAVENTLSKRAADTLKNAVK 179  
DB 120 QSSHQLGVAGGATDEAKKALLRTNAIKDGADELKLFKSVEISLAKAADALANSVN 179  
QY 180 ELTSPV 186  
DB 180 ELTSPV 186

RESULT 10  
S70284  
outer surface protein C - Lyme disease spirochete  
C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70284  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in *OspC* variation in Lyme disease  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70284  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-191 <LIV>  
A:Cross-references: EMBL:LA2896; NID:9858724; PIDN:AAB37004.1; PID:91695221  
A:Experimental source: strain 27579  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 34.2%; Score 614; DB 2; Length 191;  
Best Local Similarity 69.6%; Pred. No. 3.3e-22;  
Matches 128; Conservative 22; Mismatches 34; Indels 0; Gaps 0;

QY 3 CANSKGDNANASANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKAIG 62  
DB 1 CANSKGDNNTSANSANSADESVKGNPLTEISKRTTESNAVAVLAKVEVTLASIDELATKAIG 60  
QY 63 KKGNNGLNANQSKNTSLSGAVASDLIAEKLVNLKNEELKRIIDTAKOCSTFTNKL 122  
DB 61 OKIDANGGVAGANGSLAGAVASFLITIKLSALNSDEKLVAKVKKCSDFTKLK 120  
QY 123 SEHAVLGLDNLTDNAORAILKKHANKDGAELKLFKAVENTLSKRAADTLKNAVKELT 182  
DB 121 NGNQLGLAATDNNAKAAILKTKGTNDKGAELKLDSDVESLVKAAQVMLTNSVKELT 180  
QY 183 SPV 186  
DB 181 SPV 184

RESULT 11  
S69922  
outer surface protein C precursor - *Borrelia afzelii* (strain PLJ7)  
C:Species: *Borrelia afzelii*  
A:Variety: strain PLJ7  
C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69922; S72675  
R:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, H.  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia*  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69922  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-212 <JAU>  
A:Cross-references: EMBL:X81523



## RESULT 15

I40144

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000

C:Accession: I40144; S70282

R:Stevenson, B.; Barthold, S.W.

FEMS Microbiol. Lett. 124, 367-372, 1994

A:Title: Expression and sequence of outer surface protein C among North American isolate

A:Reference number: I40143; MUID:95154673

A:Accession: I40144

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 &lt;RES&gt;

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70282

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 19-210 &lt;LIV&gt;

A:Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218

A:Experimental source: strain 297

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 33.6%; Score 603.5; DB 2; Length 210;

Best Local Similarity 68.6%; Pred. No. 1.1e-21;

Matches 129; Conservative 26; Mismatches 30; Indels 3; Gaps 3;

QY 1 MACNNSGKDGANASANSADSVKGNLITETISKITESNAVLAKEVEETLLASIDELATKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNLITETISKITESNAVLAKEVEETLLASIDELA-KA 75

QY 61 IGGKIGNN-GLEANSKNTSLSGAVATSDLAETKLVLYKNE-ELKEKIDTAKOCSTEEFT 118

Db 76 IGGKIRKNOVSILNEADHNGSLISGAVLISTLTKKISATKDSGELKAETEKAKKCSSEFT 135

QY 119 NKLKSEHAVLIGDNTDDNAORAIILKHKANKDKGAEELEKIFKAVENLSKAQDTLKNAV 178

Db 136 AKLKEGHEHDLGKGYTDNAKAKIILKTNNDKGADLEKLEFESVKNLSKAKEMLTNSV 195

QY 179 KELTSPIV 186

Db 196 KELTSPV 203

Search completed: March 18, 2002, 09:56:48  
Job time: 371 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:52 ; Search time 39.62 Seconds  
(without alignments)  
340.552 Million cell updates/sec

Title: US-09-596-746A-38  
Perfect score: 1797

Sequence: 1 MACNNSGKQDNASANSNADES.....ENLSKAQDTLKNAVKELTS 368

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	623.5	34.7	210	1	OSCL_BORBU
2	588.5	32.7	212	1	OSCL_BORBU
3	346	19.3	214	1	VM24_BORHE
4	311.5	17.3	215	1	VM03_BORHE
5	172.5	9.6	1790	1	USO1_YEAST
6	166.5	9.3	944	1	NUP1_YEAST
7	163.5	9.1	1167	1	CAGA_HELPF
8	162.5	9.0	2017	1	MYSN_DROME
9	159.5	8.9	1875	1	MLP1_YEAST
10	158.5	8.8	705	1	YNP9_GAEEL
11	157.5	8.8	1251	1	RBP2_PLAVB
12	157	8.7	1639	1	LMG1_DROME
13	156.5	8.7	1013	1	AGOD_DROME
14	154	8.6	1526	1	MYSD_SCHPO
15	153	8.5	548	1	CEAK_ECOLI
16	152.5	8.5	1391	1	MST2_DROHY
17	151.5	8.4	1966	1	MYSB_GAEEL
18	150	8.3	1147	1	CGAL_HELPF
19	149.5	8.3	1147	1	LMB1_HUMAN
20	149.5	8.3	2116	1	MYSD_DICDI
21	148.5	8.3	1957	1	YDB6_SCHPO
22	148.5	8.3	2349	1	TPR_HUMAN
23	148	8.2	1433	1	REST_CHICK
24	148	8.2	1701	1	MSP1_PLAFR
25	147.5	8.2	1102	1	MYSC_CHICK
26	147.5	8.2	1182	1	CGA2_HELPF
27	147	8.2	845	1	SCP1_MESAU
28	147	8.2	918	1	YMB_CAEEL
29	147	8.2	1427	1	REST_HUMAN
30	146.5	8.2	1637	1	MRSP_STAU
31	146.5	8.2	1786	1	LMB1_MOUSE
32	145.5	8.1	1938	1	MYH4_RABIT
33	145	8.1	539	1	MYH3_HYDAT

34	144.5	8.0	1044	1	YAF3_SCHPO	00857 schizosacch
35	144	8.0	483	1	M6_STRPY	P08089 streptococ
36	143.5	8.0	1509	1	MYSN_ACACA	P05659 acanthamoeb
37	143	8.0	364	1	VM21_BORHE	P21875 borrelia he
38	143	8.0	1186	1	CAGA_HELPF	P55980 helicobacte
39	143	8.0	1727	1	ALM1_SCHPO	09uek5 schizosacch
40	142.5	7.9	2230	1	GOG4_HUMAN	013434 homo sapien
41	142	7.9	522	1	CEAL_HUMAN	P02978 escherichia
42	142	7.9	1186	1	SMC_BACSU	P51834 bacillus su
43	141.5	7.9	564	1	M12_STRPY	P19401 streptococ
44	141	7.8	276	1	SFAS_CHLRE	039618 chlamydomon
45	141	7.8	1701	1	MSP1_PLAFM	P08569 plasmodium

## ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD	PRT	210 AA.
AC	007337				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB819.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OG	Plasmid lp54.				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=9329332; PubMed=8098841;				
RA	Jauris-Heipke S., Fuchs R., Moiz M., Preac-Mursic V., Schwab E.,				
RA	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=9329332; PubMed=8478108;				
RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant				
RT	major outer surface protein of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=94041630; PubMed=8225587;				
RA	Padua S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=96025162; PubMed=7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.";				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=98065943; PubMed=9403685;				
RA	Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

RA Uterback T., Matthey L., McDonald L., Artlach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, *Borrelia*  
 RT burgdorferi".  
 RL Nature 380:580-586(1997).  
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X69596; CAA49306.1; -  
 EMBL: U01894; AAL6058.1; -  
 EMBL: D49497; BAA08457.1; -  
 EMBL: AE000792; AAC66329.1; -  
 TIGR: BB19; -  
 InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF001441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 KW Complete proteome.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 210 OUTER SURFACE PROTEIN C.  
 FT LIPID 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 SQ SEQUENCE 210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match 34.7%; Score 623.5; DB 1; Length 210;  
 Best Local Similarity 69.5%; Pred. No. 5.3e-22;  
 Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 1 MACNNSGKDGNAANSADSVKGNLTETSKRTESNAVLAVKVEETLLASIDELATKA 60  
 DB 17 ISCNNSGKGNTSANSADSVKGNLTETSKRTESNAVLAVKVEETLLASIDELATKA 76  
 QY 61 ICKKI-GNNGLEANSKNTSLSGAIAISDLIAEKLNVKNELEKIDTAKCSEFTTN 119  
 DB 77 ICKKHONNGDLTENNNHSLLAGAIAISTLIKOKLDIGKNGLEKIDAKKCEFTTN 136  
 120 KIKSEHAVLGLDNLTDNAORAILKHKANKDGAELKFKAVENSKAODTLKNAVK 179  
 DB 137 TKLKEHTDLGKRGVTDADAKKAILKTNGTKRGAELEKLESEVYLSKAKKEMLANSVK 196  
 QY 180 ELTSPV 186  
 DB 197 ELTSPV 203

RESULT 2  
 OSC2\_BORBU STANDARD; PRT; 212 AA.  
 AC Q08137;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE OUTER SURFACE PROTEIN C PRECURSOR (PC).  
 GN OSPC.  
 OS *Borrelia burgdorferi* (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
 OX NCBI\_TaxID:139;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=PRO;  
 RX MEDLINE=92219995; PubMed=1560779;  
 RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

RA Soutschek E.;  
 RT "Molecular analysis and expression of a *Borrelia burgdorferi* gene  
 RT encoding a 22 kDa protein (PC) in *Escherichia coli*.";  
 RL Mol. Microbiol. 6:503-509(1992).  
 RN [2]  
 RP SEQUENCE OF 1-205 FROM N.A.

RC STRAIN=DK26;  
 RA MEDLINE=94075528; PubMed=8253951;  
 RX Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
 RT "Polymorphism in ospC gene of *Borrelia burgdorferi* and  
 RT immunoreactivity of OspC protein: implications for taxonomy and for  
 RT use of OspC protein as a diagnostic antigen.";  
 RL J. Clin. Microbiol. 31:2570-2576(1993).  
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X62162; CAA44093.1; -  
 EMBL: X73624; CAA52003.1; -  
 InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF001441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 212 OUTER SURFACE PROTEIN C.  
 FT LIPID 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 SQ SEQUENCE 212 AA; 22499 MW; C206C231FBE2E7D4 CRC64;

Query Match 32.7%; Score 588.5; DB 1; Length 212;  
 Best Local Similarity 68.8%; Pred. No. 1.9e-20;  
 Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNNSGKDGNAANSADSVKGNLTETSKRTESNAVLAVKVEETLLASIDELATKA 59  
 DB 17 ISCNNSGKGDSASTNPADDSAGPNLTETSKRTESNAVLAVKVEETLLASIDELATKA 76  
 QY 60 AIGKKI-GNNGLEANSKNTSLSGAIAISDLIAEKLNVKN-BELKEKIDTAKCSEFT 117  
 DB 77 AIGKIDNNNGSLAALNNNGSLLAGAIAISTLIKSLKNELEKIDTAKKACSEFT 136  
 118 TNKISEHAVLGLDNLTDNAORAILKHKANKDGAELKFKAVENSKAODTLKNA 177  
 DB 137 TNKLSGHADLGKODATDDAKAAILKTHTTKGAKERDFEVEGLKKAQVALTNS 196  
 QY 178 VKELTSPV 186  
 DB 197 VKELTSPV 205

RESULT 3  
 VM24\_BORHE STANDARD; PRT; 214 AA.  
 AC P32778;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.  
 GN VMP24.  
 OS *Borrelia hermsli*.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
 OX NCBI\_TaxID:140;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```

RC STRAIN=SSP, HSI SEROTYPE 24;
RX MEDLINE=9313110; PubMed=1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids
are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL, L04786; AAA22964.1; -
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
FT SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;
SQ

```

```

Query Match 19.3%; Score 346; DB 1; Length 214;
Best Local Similarity 42.1%; Pred. No. 1.2e-09;
Matches 82; Conservative 40; Mismatches 57; Indels 16; Gaps 5;

```

```

OY 1 MACNSGKQGNASANSDESVKGP-----NLTEISKITTESNAVAVLAKVEVETLLASIDEL 56
DB 17 MSCNNGPPE-----LKSDEVAKSDGTVLDLAKYSKIKKSAFSAVKEVETLVKSYDEL 71
OY 57 ATKAIGKTKIGN-NLEPANSKNTSLSGAVALSDIAEKLVK-----NEELKEKIDTA 110
DB 72 A-KATKRIKKNBGLDTTEAGQNGSLAGVHSYSAVKIKVGALETTSGISNELKRTITEV 130
OY 111 KOCSTFTKRLKSEHVAVLGLDNLTDNAORAILKHKANKDKGALEKLEKFAVENISKA 170
DB 131 KSKAEAFLLMKDKGHELGKADSDDTTKAIKKKNSDKTKGASSELALNTAVDALILKAA 190
OY 171 QDTLKAIVKELTSP 185
DB 191 EGEVEAIIKELTAPV 205

```

RESULT 4  
VMP3\_BORNE  
ID VMP3\_BORNE STANDARD; PRT; 215 AA.  
AC Q02448;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.  
GN VMP3.  
OS Borrelia hermsli.  
OC Plasmid.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=140;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SSP, HSI SEROTYPE 3;  
RX MEDLINE=9313110; PubMed=1484486;  
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

```

RT are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP24.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL, L04789; AAA22967.1; -
InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
FT SEQUENCE 215 AA; 23139 MW; 684C74D35F87C771 CRC64;
SQ

```

```

Query Match 17.3%; Score 311.5; DB 1; Length 215;
Best Local Similarity 39.8%; Pred. No. 4e-08;
Matches 78; Conservative 36; Mismatches 65; Indels 17; Gaps 5;

```

```

OY 1 MACNSGKQGNASANSDESVKGP-----NLTEISKITTESNAVAVLAKVEVETLLASIDEL 56
DB 17 MSCNNGPPE-----LKSDEVAKSDGTVLDLAKYSKIKKSAFSAVKEVETLVKSYDEL 71
OY 57 ATKAIGKTKIGN-NLEPANSKNTSLSGAVALSDIAEKLVK-----NEELKEKIDT 109
DB 72 A-KATKRIKKNBGLDTTEAGQNGSLAGVHSYSAVKIKVGALETTSGISNELKRTITEV 130
OY 110 AKOCSTFTKRLKSEHVAVLGLDNLTDNAORAILKHKANKDKGALEKLEKFAVENISKA 169
DB 131 VKKSEAFVYQVSKHIDLAKEGVTDAAHAKSAILVTGTRDKGAELIKLTAIDELKA 190
OY 170 AODTLKAIVKELTSP 185
DB 191 ANDAVETVIKELTASV 206

```

RESULT 5  
USOL\_YEAST  
ID USOL\_YEAST STANDARD; PRT; 1790 AA.  
AC P25386;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.  
GN USOL OR INT1 OR YDL058W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=X2180-1A;  
RX MEDLINE=91185402; PubMed=2010462;  
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
Yamasaki M.;  
RT "A cytoskeleton-related gene, usol, is required for intracellular  
protein transport in Saccharomyces cerevisiae.";  
RL J. Cell Biol. 113:245-260(1991).  
RN [2]

RP SEQUENCE OF 782-1790 FROM N.A.  
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,  
 RA Kendrick K.E.;  
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -1- DOMAIN: THE RODLIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X54378; CA38253.1; -  
 DR EMBL: L03168; AAB00143.1; -  
 DR EMBL: U53668; AAB6659.1; -  
 DR PIR: A38455; A38455.  
 DR HSP: P80220; IDIP.  
 DR SGD: S0002216; USOL.  
 DR InterPro: IPR002017; Spectrin.  
 KM Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
 FT DOMAIN 1 724  
 FT DOMAIN 725 1790  
 FT DOMAIN 465 487  
 FT DOMAIN 991 1790  
 FT DOMAIN 1172 1786  
 FT CONFLICT 847 847  
 FT CONFLICT 924 924  
 FT CONFLICT 1253 1253  
 FT CONFLICT 1319 1319  
 FT CONFLICT 1461 1461  
 FT CONFLICT 1581 1581  
 FT CONFLICT 1600 1600  
 FT CONFLICT 1661 1661  
 FT CONFLICT 1772 1772  
 SC SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FDA818 CRC64;  
 Query Match 9.6%; Score 172.5; DB 1; Length 1790;  
 Best Local Similarity 22.1%; Pred. No. 0.6;  
 Matches 103; Conservative 66; Mismatches 152; Indels 145; Gaps 20;  
 OY 3 CUNSCKDG-----NASANSDESIVKPNLFEISKI-----TESNAVVLAVK 44  
 DB 929 CUNLSKEKEHISKELVEKSRFQSHDNV--AKLEFKLSANNYKMOENESLIRAVE 986  
 OY 45 E-----VETLASIDELATKAIKKIGNGNGLAN--OSKNTSLSGAVASDLIAEK 94  
 DB 987 ESKNESSIQLSNLONKIDMSOEKENFQIERGSIETKNEIQLKLT-----ISDLQTK 1038  
 OY 95 LNVLKNE-----LKEKIDAKOOSTEFTNKLMS-----EHAVALGLDNL- 133  
 DB 1039 EELISKSSKDEYESQISLKEKLETRATTANDNVNKKISELTKTREELAEALAAVYK 1098  
 OY 134 -----TDDNAORAI-----LKHANKDK-----AALEKLFKAVENTLS 167  
 DB 1099 NELETKELETSEKALKEVKNENHLKEKIDLEKATETTKQALNSLRANLESLEKEHEDLA 1158  
 OY 168 ---KAADDTLKNVAKELTSPYVHGNNSRKCGNASTNSADESVKPNLFEI-----SKKIT 219

DB 1159 AQLKTEEDIANKEROYNEIQLN-----DEITSTQENESIKKRN-DELEGEVAMKST 1213  
 OY 220 ESNAVVLAVKEVELLASIDELATKAIKKIGNGNGLANOSKNTSLSGAVASDLIAEK 279  
 DB 1214 SEEOSNLKKESEIDALNLQIKELKK-----NETNEAELLESINSVES----- 1255  
 OY 280 LNVLKNEELKEKIDAKOOSTEFTNKLSEHAVALGLDNLTDNNAORAIKKHANK----- 334  
 DB 1256 -ETVKIKELQDCNFKREKVESLEDKLASE-----DKNSKYLLEKRESEKIKEL 1305  
 OY 335 DKGAELEKLFKAVENTLSKA-----AQDTLKNVAKEL 366  
 DB 1306 DAKTELKIQLEKITNLKSAKESSELSRLKTTSSERKNAEOL 1351  
 RESULT 6  
 NOFL\_YEAST STANDARD; PRT; 944 AA.  
 AC P32380;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE NUPI PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).  
 GN NUPI OR SPC110 OR YDR356W OR D9476.3.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=92176232; PubMed=1541631;  
 RA Mirzayan C., Copeland C.S., Snyder M.;  
 RT "The NUPI gene encodes an essential coiled-coil related protein that  
 RT is a potential component of the yeast nucleoskeleton.";  
 RL J. Cell Biol. 116:1319-1332(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94064779; PubMed=7503995;  
 RA Klimartin J.V., Dyos S.L., Kershaw D., Finch J.T.;  
 RT "A spacer protein in the Saccharomyces cerevisiae spindle poly body  
 RT whose transcript is cell cycle-regulated.";  
 RL J. Cell Biol. 123:1175-1184(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
 RA Favellio A., Fulton L., Gattung S., Greco T., Kirsten J.,  
 RA Kucada T., Hallsworth K., Hawkins J., Hillier L., Jier M.,  
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,  
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,  
 RA Rifken L., Riles L., Tatch A., Trevisan E., Vignati D.,  
 RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL  
 CC ROLE IN CROSS-LINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT  
 CC IS ESSENTIAL FOR GROWTH.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE  
 CC NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE  
 CC NUCLEOLUS.  
 CC -1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: Z11582; CA57668.1; -  
 DR EMBL: X73297; CA51733.1; -  
 DR EMBL: U28372; AAB64791.1; -

DR PIR: S26710; S26710.  
 DR PIR: S34288; S34288.  
 DR SGD: S0002764; NUP1.  
 KW Coiled coil; Nuclear protein; Phosphorylation.  
 FT DOMAIN 164 791 COILED COIL.  
 FT DOMAIN 54 59 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 726 731 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 742 747 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 731 944 ARG/TYR-RICH.  
 SQ SEQUENCE 944 AA; 111781 MW; 04FAA074BB8A0BC8 CRC64;

Query Match 9.3%; Score 166.5; DB 1; Length 944;  
 Best Local Similarity 23.4%; Pred. No. 0.55;  
 Matches 100; Conservative 71; Mismatches 168; Indels 89; Gaps 22;

6 SGKGNASANSADSVKGNLPIEISKITESNAV-VLAKEVEETLLASIDE-LATKAIGK 63  
 91 SPRSGN-----VDSKRRRLIDDLKKDVPMSQPLKEQEVREHQMKKRFRALSKILGK 145  
 64 K-----IGNGLEANSKNTSLSGAVASIDLAIEKLVNKN-ELKEKIDTAKQS 114  
 146 RHITVANSISNKELYINEIKSLK-----HEIKELREKNDTNNYDTLEEDTDLKRL 200  
 115 TEFTNKLSEHAVLGIDNLTDDNAORAILKKHANKDGALEKLFKAV-----ENLS 167  
 201 QALEKELDAKNKIVN-SRKVDHS--GCIERROMERKLELEKRLKTVADQVLELNN 257  
 168 -----KAAQDTLKNVAVKELTSPVHGNSRRDGNAS-----TNSADE-SVKGPN 210  
 258 DVQSLKRSKEDKLNLNMLNLE--LKSNAEKQDTLEFKKNELKRTNELKIKIDE 315  
 211 L-TEISKITESNAVVLAKVEETLL-----ASIDELATKAIGKGNNGLEAN-OSKN 262  
 316 MDQLQKOKESKRLKDELNETKPESENGSOSAKENELMKLNKTAE--LEEEISTKN 373  
 263 TSLSGAVASIDLIAE-----KLVN-----LKNELKEKIDPAKOCSTFTNKLKSEH 310  
 374 SOLIAKEGKLASLMAQITOLESKLNORDSOLGSRPEELKATNDKLOK-----DIRIR 428  
 311 AVLGIDNLTDD-----DNAORAILKKHANK-----DKGALELEKLFKAVENLSKAAQDT 358  
 429 TVSKDEIIDLOKKVKVQLENDLFIYKKTTHSEKTIYTNELSKDKLIKLENDLKVAKQEK 488  
 359 LKNVAVKEL 366  
 489 YSKMEKEL 496

## RESULT 7

BLHELPJ STANDARD; PRT; 1167 AA.

AC 092LT1;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)  
 DE (CAG PATHOGENICITY ISLAND PROTEIN 26).  
 GN CAGA OR CAI OR CAG26 OR JHP0495.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 NCBI\_TaxID=85963;  
 RN NCBI\_TaxID=85963;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99120557; PubMed=9923682;  
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
 RA Tummiano P.D., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 gastric pathogen Helicobacter pylori.";

RL Nature 397:176-180(1999).  
 CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,  
 CC OR FUNCTION OF THE CYTOTOXIN.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: A601483; AAD06073.1; -  
 KW Antigen; Complete proteome.  
 FT DOMAIN 246 249 POLY-THR.  
 FT DOMAIN 882 889 POLY-ASN.  
 SQ SEQUENCE 1167 AA; 129729 MW; FD5E8B81CEBD0F2 CRC64;

Query Match 9.1%; Score 163.5; DB 1; Length 1167;  
 Best Local Similarity 26.2%; Pred. No. 0.95;  
 Matches 102; Conservative 52; Mismatches 127; Indels 109; Gaps 20;

24 PNLTETIS-----KKTESNAVVLAKVEETLLASIDELATKAIGKGNNGLEA 72  
 545 PNLNLAITSVVRDDELKILAKLSQPEANKLVKDFLSKKELVKAL-----NPNKAV 599  
 73 NQSKNTSLSGAV-----AISDLIAEKLVNKNELKEKIDTAKOCSTFTNKLKSEHAV 127  
 600 AEAKNT-----GNVDEVKQAKDL--EK--SLKKREPLEK-DVANKLSKSGNKKMEAK- 649  
 128 LGIDNLTDDNAQR-----AIIKKHANKDGALEKLFKAVENLSKAAQDTLKNVKEKLS 183  
 650 -----SQANSQKDEIFALINKKANRDARAI-----AAQNLKGIKRELSD 689  
 184 PIVHGNSRRDGNASTNSADESVKGNLPIEISKITESNAVVLAKVE-----VET 233  
 690 KLENINNDLKD---FSKSPDEFKNGKN-KDPSKABETLAKLGSVKDGIPEMISYKN 745  
 234 LLASIDELATKAIGKIGNNGLEANSKNTSLSGAVASIDLIAEKLVNKNELKEKID 293  
 746 LNALNEF-----KNGKNKDFSKVQ-----AKSDLENSIKYIINORITKQVD 789  
 294 TAKQ-----CSTFTNKLKSEHAVLGIDNLTDDN-AQRALKKHANKDGALEKLFK 346  
 790 NLNDAVSVAATGDFSG---VEQALADLKNSKQLAQAOAKNEDFNTGNKSA---LYQ 842  
 347 AVEN-----LSKRAQDTLKNVAVKEL 366  
 843 SVKNGVNGTLVNGLSKRAEATTLKSNFSDI 872

## RESULT 8

MYSN.DROME STANDARD; PRT; 2017 AA.

AC 099323;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DE MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).  
 GN MYOSIN  
 OS Drosophila melanogaster (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN NCBI\_TaxID=7227;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90349606; PubMed=2117279;  
 RA Ketchum A.S., Stewart C.T., Stewart M., Klehart D.P.,  
 RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain  
 transcript: conserved sequences in the myosin tail and differential

```

RT  splicing in the 5' untranslated sequence."
RL  Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC  -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC  CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC  -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC  SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC  -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC  -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL: M5012; AAA28713.1; -.
PIR: A36014; A36014.
DR  HSSP: P08799; 1MD.
DR  FLYBASE: FBgn0005634; zip.
DR  InterPro: IPR000048; IQ.
DR  InterPro: IPR002928; Myosin_tail.
DR  InterPro: IPR001609; myosin_head.
DR  Pfam: PF00612; IQ; 1.
DR  Pfam: PF00063; myosin_head; 1.
DR  Pfam: PF01576; Myosin_tail; 1.
DR  PRINTS: PR00193; MYOSINHEAVY.
DR  PRODOM: PD000355; myosin_head; 1.
DR  SMART: SM00015; IQ; 1.
DR  SMART: SM00242; MISC; 1.
DR  SMART: P55096; IQ; 1.
KW  Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW  ATP-binding; Calmodulin-binding.
FT  DOMAIN 1 829 859 10 COILED COIL (POTENTIAL).
FT  DOMAIN 886 2017 232 ATP.
FT  NP_BIND 225 232 250 KDA/50 KDA JUNCTION.
FT  DOMAIN 250 260 50 KDA/20 KDA JUNCTION.
FT  DOMAIN 682 694 727 ACTIN-BINDING.
FT  DOMAIN 705 727 742 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT  DOMAIN 742 758 1303 2017 LIGHT MEROMYOSIN (LMN).
FT  DOMAIN 1303 1970 1971 2017 ALPHA-HELICAL TAILPIECE (LMN).
FT  DOMAIN 1971 2017 GLOBULAR TAILPIECE.
FT  VASPLIC 1 45 MISSING (IN SHORT ISOFORM).
FT  SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;
-----
Query Match 9.0%; Score 162.5; DB 1; Length 2017;
Best Local Similarity 23.8%; Pred. NO. 1.9;
Matches 96; Conservative 63; Mismatches 139; Indels 101; Gaps 18;
-----
QY 26 LTESKRTTESNAVY-----LAVEVE---TLASIDELTKAIGKIG----- 66
DB 1108 VADLEQENRRVQYDEMQLAKKEELLTQTLRIDESKTKATAQAQELSSQALAI 1167
QY 67 NNGLEANSKRTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQCKSTFTNKLKSEHA 126
DB 1168 QEDLEAEKAKARAK-----AEKYRDLSELLEALKNELDLSDTTAAQ--QELRSKREGELA 1221
QY 127 VLGLDNLTDNAORAIILKHAH-----KDKGAELKELFKAVENISKA-----AQD 172
DB 1222 TL-----KKSLEETVNHHCYVLADMRKHQSDELNSINDOLENRKAKTVLEKAKG 1271
QY 173 TLKNAVELSPIVHGNSNRKGNASTNSADESVKGNLTLEISKRTTESNAVLAVK--E 230
DB 1272 TLEENMDLATELRSVNSRQENDR-----RRKQAESQALQVLAEL 1314
QY 231 VETLASIDELATAIKGKIGN--NGLEANSKRTSLSGAYAISDLIAEKLNVKNEEL 288
DB 1315 IERARSELQEKCTK-LOOEENITNLEAEELKASAAVKSASNMESOLTEAOGLLE-EET 1372
= 3

```

```

QY 289 KEKIDTAQCKSTFTNKLKSEHAVLGLDNLTDNAORAI-----LKRHAKND 335
DB 1373 KGLIGLSK-----LROISEKEALQOELEDEDAKRNVERKLAETVTQMGKIKKAED 1427
QY 336 KG-AALE-----KLFAVENISK-----AADDTLKNAVELTS 368
DB 1428 ADLAKLEEGKKRLNDIDALERQYKELIAQNDRLDKSRKKIQS 1471
-----
RESULT 9
ID  MLPI_YEAST STANDARD: PRT: 1875 AA.
AC  002455;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  20-AUG-2001 (Rel. 40, Last annotation update)
DE  MYOSIN-LIKE PROTEIN MLPI.
GN  MLPI OR YKR095W OR YKRA15.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S288C.
RX  MEDLINE=93247549; PubMed=8483450;
RA  Koelling R., Nguyen T., Chen E.Y., Botstein D.;
RT  "A new yeast gene with a myosin-like heptad repeat structure.";
RL  Mol. Gen. Genet. 237:359-369(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94205265; PubMed=8154186;
RA  Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA  Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT  "The complete sequence of a 15,820 bp segment of Saccharomyces
RT  cerevisiae chromosome XI contains the UBI2 and MPl1 genes and three
RL  new open reading frames.";
RL  Yeast 9:1349-1354(1993).
CC  -1- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
CC  REPAIR.
CC  -1- SIMILARITY: SOME, TO THE TPR ONCOGENE.
CC  -1- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPl1".
-----
QY 26 LTESKRTTESNAVYLAVENTLTLASIDELATAIKGKIGNNGLEANSKRTSLSGAY 85
DB 647 ISQITRESTEN--MSLLNNEIDLYDKSDISIK-LGKRSRIIALEEFK---LLSNLT 700
-----
Query Match 8.9%; Score 159.5; DB 1; Length 1875;
Best Local Similarity 21.3%; Pred. NO. 2.4;
Matches 96; Conservative 80; Mismatches 155; Indels 119; Gaps 18;
-----

```

QY 86 -----AISDLIAEKLNVLKNEELKEKIDTAQCSSTFTNKLSEHAVLGLDNLTDNAORA 141  
 DB 701 DLTAKENLDOLRRKFDLQNTILKQDSKTHETLNEYVSCSKSIYETELLNLKEEOKLR 760  
 QY 142 ILKHKANKDKGAEELEK-----LFKAVENLSKAAODTLKNAVELSPYVGNNSKDCN 196  
 DB 761 HEKMLKQGLNKLSPEKSDLRIMVQLQTLQKEREDLLETRKSCCKIDELDELAJELK 820  
 QY 197 ASTNSADESVKPNLTERISKITE--SNAVVLAVKEVEFTLLASIDELAT-----KA 245  
 DB 821 KETSDCKDHHRK--OLEEDNNSNIEMTYQNKTEALKDYESVITSVDSKQTDIEKLYQYKVS 878  
 QY 246 IGKK-----IGNNGLEANSKNTSLSGAA----- 271  
 DB 879 LEKELEBEDKIRLHTYVWDETINDDSLKELEKSKNTLTDASQIKYEDLYETTSQISQ 938  
 QY 272 -----ISDL-----IAEKLNVK-----NEEL-----KERIDPA 295  
 DB 939 OTNSKLDSEFKDFTNQIKNLTDDEKTSLEDKISLKEQMFNLNELLQKKMEKEKADFR 998  
 QY 296 KCCS-TEFTNK-----LKSEHAVLGLDNLTDNAORAI-----LKKHANKDKGA 338  
 DB 999 KRISLQNNKKEVAVKSEYE-SKLSKIQNDLDOQTIVANTQNNYEOELQKHADVSKIT 1057  
 QY 339 AELEK---LFKAVENLSKAAODTLKNAVE 365  
 DB 1058 SELREQLHTYKGOVKTLLNSRDQLENALKE 1087.

RESULT 10  
 YNP9\_CAEEL STANDARD: PRT: 705 AA.

AC P34562;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL 80.6 KDA PROTEIN T0565.9 IN CHROMOSOME III.  
 GN T0565.9  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;

RA [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Lathelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Sims M., Smailson N., Smith A., Smith K., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Watson R., Watson A., Welstock L., Wilkinson-Sproat J.,  
 RA Wohldman P.;  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RT Nature 368:32-38(1994).  
 RN [2]  
 RP REVISIONS.

RA Durbin R.;  
 CC This SWISS-PROT entry to the EMBL/Genbank/DBA databases.  
 CC -----  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL: Z27079; CAA81596.1; -  
 DR PIR: S41009; S41009  
 DR WormPep: T0565.9; CE21153.  
 DR InterPro: IPR000237; GRIP.  
 DR Pfam: PF01465; GRIP; 1.  
 KW Hypothetical protein; Coiled coil.  
 FT DOMAIN 75 137 COILED COIL (POTENTIAL).  
 FT DOMAIN 160 509 COILED COIL (POTENTIAL).  
 FT DOMAIN 562 641 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 705 AA; 80637 MW; 9E89873F5FC04566 CRC64;

Query Match 8.8%; Score 158.5; DB 1; Length 705;  
 Best Local Similarity 25.0%; Pred. No. 0.91;  
 Matches 91; Conservative .52; Mismatches 140; Indels 81; Gaps 14;

QY 25 NLTEISKRTTESNAVVLAVKEVEFTLLASIDELATKAIGKRTGNGLSEANSKNTSLSGA 84  
 DB 94 NOTALMEKTLKAKMSLEQVKKDAENLKLINKELTTESAKVENNPFC-----TECLSKS 148  
 QY 85 YAISDLIAEKLNVLKNEELKEKIDTAQCSSTFTNKLSEHAVLGLDNLTDNAORAIK 144  
 DB 149 GALIELEKEVYF-----EMKEK-----ATRA--- 168  
 QY 145 KHANKDKGAEELEKLFKAVENLSKAAODTLKNAVE---LTSPIVHGNNRKDGNASTNS 201  
 DB 169 -----DISLELDLSKSVQNLNLRDKTEALIKAOEVTENDLEVNNKKEKN-NTKS 222  
 QY 202 ADESVKPNLTERISKITESNAVVLAVKEVEFTLLAS---IDELATKAIGKRTGNGLSEA 257  
 DB 223 SIEKLTREN-TRLTALQODEK---IKSADFEARLRSAECHRIVELSDQGNKGLAKKMA 278  
 QY 258 NQSKNTSLSGAVAYISDLIAEKLNVLKNEELKEKIDTAQCSSTFTNKLSEHAVLGLDN 317  
 DB 279 ESENRCGILTEEAVDVLSKSENEKL-LAKNEEFSAKLVSSEK--EFAEFKKRSHPELEKKG 334  
 QY 318 LTPDNMORALKKHANKDKGAEELEK-----FKAVENLSKA-----AODTLKNV 363  
 DB 335 KOEDETRKALELEKSK-VYITTELEQADOTROEHFPTVEDLASSRDKAEERLEKTVLK 393  
 QY 364 KEIT 367  
 DB 394 SEIT 397

RESULT 11  
 RBP2\_PLAVB STANDARD: PRT: 1251 AA.  
 ID RBP2\_PLAVB  
 AC Q00759;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).  
 GN RBP2.  
 OS Plasmodium vivax (strain Belem).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxId=31273;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92315338; PubMed=1617731;  
 RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
 RT "A reticulocyte-binding protein complex of Plasmodium vivax  
 RT merozoites.";  
 RL Cell 69:1213-1226(1992).  
 CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
 CC HUMAN RETICULOCYTE CELLS.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M88098; AAA29744.1; -  
KW Malaria; Receptor; Membrane.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CNC64;

Query Match 8.88; Score 157.5; DB 1; Length 1251;  
Best Local Similarity 20.08; Pred. No. 1.9;  
Matches 106; Conservative 77; Mismatches 151; Indels 181; Gaps 21;

QY 9 DGNASANSADSESVK---GPNLEISKIKTESNAVYLA---KEVEFLASIDELAT 58  
DB 671 EGHGNVQVLENIKELVDENNLSIDLKQATGKNEELQKTHSLTKNAKKTILGHVPTSA- 729  
59 KAIGKKT-----GNNGLEA-NOSKNTSLISGAYAISDLIT 91  
DB 730 KYVGKITPPELATVELLDGAKLKTQAELKPESKNNVYLETENMSKNTNELDVHKNIDAY 789  
QY 92 AEKLVNKNELKEKIDTAQOST-----EFTNKLKS----- 123  
DB 790 KVALEILAH---SDEIDTKQKSSKLIEMGNQIYKVLINQKKNISIKSKEEAVSVK 846  
QY 124 -----EHAVLG-----LDNLTDDNACRAILKKRANKDKGAELE 157  
DB 847 IGVNSKRSKHSLSKTCSDSKSYDNIILEKQTELQNLRSNFTQE---KTMND-----SKIE 900  
QY 158 KLFKAVENTSKAODPLKNAVVELTSPYHGNNRSDGASNSADESKGPNLEISK 217  
DB 901 KI-----KIDFESLKNAKLKEGE-VNALKASSDNEHVOSKSEPV-NPALSEIEKE 950  
QY 218 ITESNAVLAKEV---ETLLASIDELATKAIGKIGN---NGLEAN----- 258  
DB 951 EFDIDSLNALDELKKTGRCESRYKLKIDVYKEISDDELINTIENNVAYIAYIKK 1010  
QY 259 -----OSKNTSLISGAYAISDLITAEKL-----NV 282  
DB 1011 NEDTVQVQVLTNEHENTKQVNSHEPTNEDKNSKEELTKAVTSKTIISKLGVITIEV 1070  
QY 283 LKNEELKEKIDTAQOSTEFTNKLKSEHVLGLDNLTDNAQRALTKHANKDKGAELE 342  
DB 1071 NENTENNTIESSAKTELEALY-NELKNKTSLSLEYITQTSNEVYKLEPMKSNADK---YIDVS 1126  
343 KLFKAVENTSKA-----AODTLKNAVKE 366  
DB 1127 KIFNTVLDTKQKSNIVTNGHSINNVDKIKGLKQL 1161

RESULT 12  
LNG1\_DROME STANDARD; PRT: 1639 AA.  
AC P15215; Q24373; Q9VT18;  
DT 01-APR-1990 (Rel. 14, Created).  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).  
GN LAMB2 OR LAMC1 OR LAMG1 OR CG3322.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
XN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CANTON-S, AND OREGON-R;  
RX MEDLINE=91299161; PubMed=1840513;  
RA Chi H.-C., Jumilaga D., Wang S.-Y., Hui C.-F.;  
RT "Structure of the Drosophila gene for the laminin B2 chain.";  
RL DNA Cell Biol. 10:451-466(1991).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON-R;  
RX MEDLINE=89109164; PubMed=2912972;  
RA Chi H.-C., Hui C.-F.;  
RT "Primary structure of the Drosophila laminin B2 chain and comparison with human, mouse, and Drosophila laminin B1 and B2 chains.";  
RL J. Biol. Chem. 264:1543-1550(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90037237; PubMed=2808533;  
RA Montell D.J., Goodman C.S.;  
RT "Drosophila laminin: sequence of B2 subunit and expression of all three subunits during embryogenesis.";  
RL J. Cell Biol. 109:2441-2453(1989).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Burton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,  
RA Brudnon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iregan C.,  
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [5]  
RP SEQUENCE OF 344-1639 FROM N.A.  
RC STRAIN=OREGON-R;  
RX MEDLINE=88303364; PubMed=3405777;  
RA Chi H.-C., Hui C.-F.;  
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";  
RL Nucleic Acids Res. 16:7205-7205(1988).  
RN [6]  
RP FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATING  
CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
CC -1- SUBCELLULAR LOCATION: BASEMENT MEMBRANES (MAJOR COMPONENT).  
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT

CC		WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC	-1-	DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC	-1-	SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC	-1-	SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
CC	-1-	SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch</a> ).
CC		or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> .
DR	EMBL; M58417;	AAA28665.1; .
DR	EMBL; M25063;	AAA28664.1; .
DR	EMBL; AE003551;	AAE50238.1; .
DR	EMBL; X07806;	CAG30665.1; .
DR	PIR; A31483;	MMEFB2.
DR	HSSP; P02468;	TITLE.
DR	Flybase; FBgn0002528;	Lamb2.
DR	InterPro; IPR000561;	EGF-like.
DR	InterPro; IPR001886;	LamNT.
DR	InterPro; IPR000034;	LaminIn_EGF.
DR	InterPro; IPR002049;	LaminIn_B.
DR	Pfam; PF00052;	LaminIn_B_1.
DR	Pfam; PF00053;	LaminIn_EGF_10.
DR	Pfam; PF00055;	LaminIn_Nterm; 1.
DR	ProDom; PD002082;	LamNT; 1.
DR	ProDom; PD003031;	LaminIn_B; 1.
DR	SMART; SM00180;	EGF_Lam; 10.
DR	SMART; SM00001;	EGF_like; 1.
DR	SMART; SM00281;	Lamb; 1.
DR	SMART; SM00136;	LamNT; 1.
DR	PROSITE; PS00022;	EGF_1; 8.
DR	PROSITE; PS01186;	EGF_2; 1.
DR	PROSITE; PS01248;	LAMININ_TYPE_EGF; 11.
KW	Glycoprotein;	Basement membrane; Extracellular matrix; Coiled coil;
KW	Laminin EGF-like domain;	Cell adhesion; Repeat; Signal.
FT	SIGNAL	1..33
FT	CHAIN	34..1639
FT	DOMAIN	34..298
FT	DOMAIN	299..523
FT	DOMAIN	299..358
FT	DOMAIN	359..413
FT	DOMAIN	414..460
FT	DOMAIN	461..513
FT	DOMAIN	514..523
FT	DOMAIN	524..709
FT	DOMAIN	710..1049
FT	DOMAIN	710..743
FT	DOMAIN	744..792
FT	DOMAIN	793..846
FT	DOMAIN	847..901
FT	DOMAIN	902..955
FT	DOMAIN	956..1003
FT	DOMAIN	1004..1049
FT	DOMAIN	1050..1609
FT	DOMAIN	1087..1109
FT	DOMAIN	1144..1247
FT	DOMAIN	1306..1627
FT	DISULFD	299..308
FT	DISULFD	301..322
FT	DISULFD	324..333
FT	DISULFD	336..356
FT	DISULFD	359..368
FT	DISULFD	361..384
FT	DISULFD	387..396
FT	DISULFD	399..411
FT	DISULFD	414..426
FT	DISULFD	416..432

Query Match	Best Local Similarity	Score 157; DB 1; Length 1639;
Matches 104; Conservative	61; Mismatches 158; Indels 108; Gaps	21
2 ACNNGKGNMNSANAD-----ESVKGNNLEISKITESNAVVLAVAEVEFTLASIDE 55	8.7%; Pred. No. 2.7;	
Db 1107 AADNSGGGQGYAEYVIDLHKHLDVR-EHLVSADPFQADANGETIDRANKTTILDQITE 1165		
Qy 56 LATKAIGKKI-----GNNGLEANOSKNTSLSGAVAISDL-----IAEKLNV-----L 98		
Db 1166 NAKKTELQALDLINDEGAQALRAKEKSYEPQGSQSIDISREARALADKLESEAQFDL 1225		
Qy 99 KN-EELKEKIDTAQKCS-----TEFTNKLKSEHAVLGLDNLTPDNOARAILKHKAND 150		
Db 1226 KNAKDAKDAVEKAHOLAKSAIDLOLKIGTELKSE--VGLF-----LSHVKOS 1270		
Qy 151 KGA---AELEKFKAVENLSKAAQDTLKAIVKVELSPIYHGNNSRKDGNASTNSADESVK 207		
Db 1271 LGTVVQTSKEARKNEEYDTAL--TLLDVNRQTOPELIDISQLKDAVAANERADDELK 1328		
Qy 208 GPNLTEISKITTESNAVVLAVAEVEFTLASIDELATKAIGKKIGNNGLEANOSKNTSLS 267		
Db 1329 --QITFELS-----NSNGELFADEFTEQEL-----TEALKR-----AEQQLEDIELE 1370		
Qy 268 GAVAYSIDLIAEKL-----NVLK-----NEELKEKIDTAQKCSFTFNN 304		
Db 1371 RAKAAHDKATKAVEGQDNTLKEANNTYEKLAFQSDPVQSSSAEKALQTVNIREKEION 1430		
Qy 305 K---LKSEHAVLGLD---NLTDNAORAILK--KHANDKAAAELEKFKAVENLSKAA 355		
Db 1431 AESLISQAEALDGNKKNAKNEAKKNAQEOQLTYADQASKD-----AELIRKKAETTKVA 1484		
Qy 356 QDTLKNAYEKL 366		
Db 1485 ARNLREPADQL 1495		



A60D\_DROME STANDARD; PRT; 1013 AA.

AC P91927; 09W160;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CALCIUM-BINDING MITOCHONDRIAL PROTEIN ANON-60DA.  
 GN ANON-60DA OR CG4589.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NX NCBI\_TaxId-7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew R.M., Basu A., Bayendale J., Beasler E.M.,  
 Beeson K.Y., Bonos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck K.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,  
 Jajuli M.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Mishina N.V., Mohanty C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 Reihart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 [12]  
 RN SEQUENCE OF 626-944 FROM N.A.  
 RP TISSUE-Ovary;  
 RC MEDLINE-99168769; PubMed-10071211;  
 Caggese C., Regone G., Perrini B., Moschetti R., de Pinto V.,  
 Calzi R., Barsanti P.;  
 "Identification of nuclear genes encoding mitochondrial proteins:  
 isolation of a collection of *D. melanogaster* cDNAs homologous to  
 sequences in the Human Gene Index database.";  
 Mol. Gen. Genet. 261:64-70(1999).  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO  
 CC FRAMESHIFTS IN POSITIONS 920 AND 930.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use.

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE003464; AAF47217.1;  
 DR EMBL: Y10912; CA71853.1; ALT\_FRAME.  
 DR FlyBase: FBgn019886; CG4598.  
 DR InterPro: IPR001813; 60S\_ribosomal.  
 DR InterPro: IPR002024; Bacterioferitin.  
 DR InterPro: IPR003871; DUF223.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR000861; REM\_repeat.  
 DR InterPro: IPR002555; RFL.  
 DR InterPro: IPR001236; 1dh.  
 DR Pfam: PF00428; 60S\_ribosomal; 1.  
 DR Pfam: PF01334; Bacteriofer; 1.  
 DR Pfam: PF02721; DUF223; 1.  
 DR Pfam: PF00036; efhand; 2.  
 DR Pfam: PF02185; HRI; 1.  
 DR Pfam: PF00056; 1dh; 1.  
 DR Pfam: PF01605; RFL; 1.  
 DR SMART: SM00054; Efh; 2.  
 DR PROSITE: PS00018; EF\_HAND; 2.  
 KW Mitochondrion; Calcium-binding; Repeat.  
 FT CA\_BIND 700 711 EF\_HAND 1 (POTENTIAL).  
 FT CA\_BIND 783 794 EF\_HAND 2 (POTENTIAL).  
 FT CONFLICT 687 687 E -> Q (IN REF. 2).  
 FT CONFLICT 690 690 K -> G (IN REF. 2).  
 FT CONFLICT 717 717 Q -> H (IN REF. 2).  
 FT CONFLICT 736 736 E -> D (IN REF. 2).  
 FT CONFLICT 740 740 K -> R (IN REF. 2).  
 FT CONFLICT 761 761 K -> E (IN REF. 2).  
 FT CONFLICT 943 943 G -> S (IN REF. 2).  
 SQ SEQUENCE 1013 AA; 113579 MW; 8C710380263F262E CRC64;  
 Query Match 8.7%; Score 156.5; DB 1; Length 1013;  
 Best Local Similarity 22.6%; Pred. No. 1.7;  
 Matches 81; Conservative 62; Mismatches 141; Indels 75; Gaps 12;  
 QY 29 ISKTTESNAVYLVAVEYETLLASIDELATYKAIGKIGNNGLEANOSKNTSLSGVAIS 88  
 DB 609 VKEPREERAALKLVRNKNMISQIDNVL-----NLEARQHQIKOAESSDYAAS 658  
 QY 89 DLAELKLVNKEELKEKIDPAKOCSTEFNTKIKSEHAVLGIDNITDQNAAILKKNAN 148  
 DB 659 SPTVPEQOMHIDELVATIRKKEASDERFRVGD-----LIVKIDAD 702  
 QY 149 KDKGAELKEKLEKAVENLSKAADTLKNAVELT-----SPIYGNNSRKDGN 196  
 DB 703 KD-GVIVNEITKAVQSIDREATNIDKKQLEPELLESLKLSRRRHEIVAIIDLMNNIK 761  
 QY 197 ASTNSADSVGPNITELSKITESTNAVYLVAVEYETLLASIDELATYKAIGKIGNNGLE 256  
 DB 762 VLKETSDEA-RLKHEALEKFEADKDGCVTVYNDIRKLES-----IGRDNK 808  
 QY 257 ANOSKNTSLSGAAYAISDLA--EKLNVKNEELKEKIDPAKOCSTEFNTKSE--HAV 312  
 DB 809 LSDK-----AIEELISLDKBOYLOAE--QAIERAIASKKEAEKLSSEVDKAD 855  
 QY 313 LGDNLITDQNAORA-----ILKKNANKDKGAELKEKLEKAVENLSKAADTLKNAVEL 366  
 DB 856 KDLKSLVNDIHDSDAKEIODIANEMRDRKEEYVPAKAKELKA-EPAFKDTATITLQMDNADL 913  
 RESULT 14  
 ID MY52\_SCHPO STANDARD; PRT; 1526 AA.  
 AC Q9US16; P78969;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)



DE MYOSIN TYPE II HEAVY CHAIN 1.  
 GN MYO2 OR SPCC645.05C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RC STRAIN-972;  
 RX MEDLINE=98075862; PubMed=9415380;  
 RA May K.M., Watts F.Z., Jones N., Hyams J.S.;  
 RT "Type II myosin involved in cytokinesis in the fission yeast,  
 RL Schizosaccharomyces pombe.";  
 RN Cell Motil. Cytoskeleton 38:385-396(1997).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA Wood V., Rajandream M.A., Barrell B.G., Rieger M.;  
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR CELL DIVISION. IT IS A COMPONENT OF THE  
 CC CDC12 'SPOT', A STRUCTURE THOUGHT TO MARK THE SITE OF SEPTATION.  
 CC MAY WORK IN CONJUNCTION WITH MYO3.  
 CC -1- SUBUNIT: BINDS TO CDC4 AND RLC1.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U75357; AAC49908.1; -;  
 CC EMBL: AL048498; CAB39901.1; -;  
 CC HSSP: P08799; 1MD.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ.1.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head.1.  
 DR SMART: SM00015; IQ.1.  
 DR SMART: SM00242; MYSC.1.  
 DR PROSITE: PS00096; IQ.1.  
 KW Myosin: Actin-binding; ATP-binding; coiled coil; Calmodulin-binding;  
 KW Alkylation.  
 PE DOMAIN 1 757 MYOSIN HEAD-LIKE.  
 PE DOMAIN 758 787 IQ.  
 FT NP\_BIND 875 1244 COILED COIL (POTENTIAL).  
 FT NP\_BIND 170 177 ATP (POTENTIAL).  
 FT DOMAIN 634 656 ACTIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 734 748 ACTIN-BINDING (BY SIMILARITY).  
 FT MOD\_RES 674 674 ALKYLATION (BY SIMILARITY).  
 FT CONFLICT 1337 1337 S -> R (IN REF. 1).  
 SQ SEQUENCE 1526 AA; 176430 MW; D71D51D6578192BA CRC64;

Query Match 8.6%; Score 154; DB 1; Length 1526;  
 Best Local Similarity 23.3%; Pred. No. 3.3; Indels 74; Gaps 18;  
 Matches 88; Conservative 76; Mismatches 139;

OY 8 KDSGANSANDESVPKPN-LTEISKRTESNAVLAVKEV-----ETLLASIDE----- 55  
 DB 842 KQNSKSEVERDLVETNNSTLAVENLLTTRATALDKEELTRKQERLANIEDSFSTKQ 901  
 OY 56 ----LATKAIGKRTGNNGLEANOSKNTSLSGAVAISDLAEKLVAKNELKEKIDYAK 111  
 DB 902 QNENLQRESASLKQINNELESELLEKTS-----KVETLLSEQ-----NELKEKLSLE 949  
 OY 112 QCSFTFTNKLKS--EHAIVGLDNLDTDDNQ--RAILKKHAKKDGAELEKLFRAVENLSK 168

DB 950 KDLDTGELSLRENNATYLVSEAEFNEOCKSLQETITVTD---AELDKLTITSPYKT 1006  
 OY 169 AADDTLNAKVELTSPYVHGNNSKRDGNASTNSADESVKGNPLTEISKRTESNAVLVA- 227  
 DB 1007 EIQE-----MRLTNQMNKSTIOQBSLS-----ESLKR-----VKLERESTLSD 1049  
 OY 228 -----VKVETLLASIDELATKAIGKRTGNNGLEANOSKNTSLSGAVAISDL-IAEKL 280  
 DB 1050 VSILKQKKEELSVLKGVOEELTNLEEKV--NYLEADVOKLPKIKKELESINDKQYQL 1107  
 OY 281 NVLKNELKEKIDTAKCSTFEFTKLKS-----EHAIVGLDNLDDNAQRAILKK-IANK 334  
 DB 1108 QATRNKELEAKV--KCC-----LNNIKSLTKELENKEEKQNLSDASLKYIELOIHEHL 1160  
 OY 335 DKGAELERLFAVENL 351  
 DB 1161 LKAVSDLENKKKEGL 1177

RESULT 15  
 CEAK\_ECOLI  
 ID CEAK\_ECOLI STANDARD; PRT; 548 AA.  
 AC 047502; P75615;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE COLICIN K.  
 GN CKA.  
 OS Escherichia coli.  
 OG Plasmid ColK-K235.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K49;  
 RX MEDLINE=96074330; PubMed=7592493;  
 RA Pilsel H., Braun V.;  
 RT "Strong function-related homology between the pore-forming colicins K  
 RT and 5.";  
 RL J. Bacteriol. 177:6973-6977(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Izard J., Chartier M., Baty D.;  
 RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF  
 CC TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING  
 CC TO DISSIPATION OF CELLULAR ENERGY.  
 CC -1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE  
 CC AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.  
 CC -1- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X87834; CAA61099.1; -;  
 CC EMBL: U27452; AAB41288.1; -;  
 DR InterPro: IPR000293; Channel\_colicin.  
 DR Pfam: PF01024; Colicin.1.  
 DR PRINTS: PR00280; CHANNELCOLICIN.  
 DR ProDom: PD002657; Channel\_colicin.1.  
 DR PROSITE: PS00276; CHANNEL\_COLICIN.1.  
 KW Antibiotic; Bacteriocin; Plasmid; Transmembrane.  
 KW TRANSMEM 505 525 POTENTIAL.  
 FT CONFLICT 155 155 S -> G (IN REF. 2).  
 SQ SEQUENCE 548 AA; 59661 MW; 2E676890080CFF6 CRC64;



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:54 ; Search time 124.19 Seconds  
(without alignments)  
433.434 Million cell updates/sec

Title: US-09-596-746A-38

Perfect score: 1797  
Sequence: 1 MACNNSCKDGNASANSADSES.....ENUSKAAQDTLKNVKELTSS 368

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.17.\*

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_plant:\*  
10: sp\_phage:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description
1	901	50.1	192	2	Q9RS52	Q9RS52 borrelia bu
2	901	50.1	209	2	Q44717	Q44717 borrelia bu
3	899	50.0	191	2	Q9S3P1	Q9S3P1 borrelia bu
4	889	49.5	192	2	O31117	O31117 borrelia bu
5	663	36.9	194	2	O31122	O31122 borrelia bu
6	661	36.8	193	2	P94234	P94234 borrelia bu
7	653.5	36.4	212	2	Q44705	Q44705 borrelia bu
8	646	35.9	202	2	O08231	O08231 borrelia bu
9	646	35.9	211	2	Q44720	Q44720 borrelia bu
10	644	35.8	201	2	Q9R0R8	Q9R0R8 borrelia bu
11	642	35.7	193	2	P94237	P94237 borrelia bu
12	639	35.6	191	2	Q44726	Q44726 borrelia bu
13	623.5	34.7	192	2	Q9R7B1	Q9R7B1 borrelia bu
14	623.5	34.7	193	2	Q9R8S3	Q9R8S3 borrelia bu
15	623	34.7	185	2	O31123	O31123 borrelia bu
16	621.5	34.6	192	2	Q9S3P3	Q9S3P3 borrelia bu
17	620	34.5	200	2	O9XDH4	O9XDH4 borrelia bu
18	619	34.4	193	2	P94236	P94236 borrelia bu
19	616	34.3	209	2	Q44883	Q44883 borrelia bu

20	614	34.2	191	2	Q9S3P0	Q9S3P0 borrelia bu
21	612	34.1	188	2	Q9XDH3	Q9XDH3 borrelia bu
22	610.5	34.0	203	2	O50620	O50620 borrelia af
23	609.5	33.9	200	2	Q9R0R9	Q9R0R9 borrelia bu
24	609.5	33.9	205	2	P96505	P96505 borrelia af
25	607.5	33.8	193	2	O31115	O31115 borrelia bu
26	607.5	33.8	201	2	P96571	P96571 borrelia ja
27	607.5	33.8	212	2	Q44670	Q44670 borrelia af
28	605.5	33.7	191	2	P70818	P70818 borrelia bu
29	605.5	33.7	194	2	P94229	P94229 borrelia bu
30	605.5	33.7	201	2	P96573	P96573 borrelia ja
31	605	33.7	178	2	O44979	O44979 borrelia bu
32	605	33.7	178	2	O44995	O44995 borrelia bu
33	603.5	33.6	210	2	O44719	O44719 borrelia bu
34	602.5	33.5	201	2	P96516	P96516 borrelia ja
35	601.5	33.5	192	2	Q9S3P2	Q9S3P2 borrelia bu
36	601.5	33.5	201	2	P96572	P96572 borrelia ja
37	600	33.4	190	2	P70819	P70819 borrelia bu
38	599	33.3	184	2	Q9S504	Q9S504 borrelia bu
39	598.5	33.3	191	2	O31120	O31120 borrelia bu
40	596.5	33.2	182	2	Q9R7B2	Q9R7B2 borrelia bu
41	596.5	33.2	212	2	O44718	O44718 borrelia bu
42	594.5	33.1	201	2	P96514	P96514 borrelia ja
43	594	33.1	181	2	O34124	O34124 borrelia bu
44	592	32.9	184	2	O34120	O34120 borrelia bu
45	591	32.9	211	2	Q9Z6C7	Q9Z6C7 borrelia va

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	192 AA.
ID	Q9RS52	Q9RS52		
AC	Q9RS52	Q9RS52		
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).			
GN	OSPC.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID-139;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-OC5;			
RX	MEDLINE-99091544; PubMed-9872945;			
RA	Wang I.N., Dykhuizen D.E., Olu W., Dunn J.J., Bosler E.M., Luft B.J.;			
RT	"Genetic diversity of ospC in a local population of Borrelia burgdorferi sensu stricto."			
RL	Genetics 151:15-30(1999).			
DR	EMBL; AF029864; AAB86547.1; "			
DR	InterPro: IPR001800; Lipoprotein_6.			
DR	pfam: PF01441; Lipoprotein_6; 1.			
FT	Prodom: PD001149; Lipoprotein_6; 1.			
FT	NON_TER 1			
FT	NON_TER 192			
SQ	SEQUENCE 192 AA; 20555 MW; FFL6D409D58C01C5 CRC64;			
Query Match	50.1%;	Score 901;	DB 2;	Length 192;
Best Local Similarity	98.9%;	Pred. No. 3;	6e-35;	
Matches 184;	Conservative	2;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1	MACNNSCKDGNASANSADSEVKGPNLFEISKTTESNAVLAWEVETLLASIDELATKA	60	
DB	7	ISCNNSCKDGNASANSADSEVKGPNLFEISKTTESNAVLAWEVETLLASIDELATKA	66	
QY	61	IGKRIKGNGLNGLANOSKNTSLSGAYAISDLIAEKLNVLNKEELKEKIDTAQKOSTEFTNK	120	
DB	67	IGKRIKGNGLNGLANOSKNTSLSGAYAISDLIAEKLNVLNKEELKEKIDTAQKOSTEFTNK	126	
QY	121	LKSEHAVLGIDNLTDDNAQAAILKKHANKDKGALELEKFKRAVENLSKAQDTLKNVKE	180	

Db 127 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 186  
 QY 181 LTSPV 186  
 Db 187 LTSPV 192

## RESULT 2

044717 PRELIMINARY: PRT: 209 AA.

AC 044717  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OSPC.  
 GN OSGC.  
 OR Borrelia burgdorferi (Lyme disease spirochete).  
 NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N40;  
 RX MEDLINE=94314484; PubMed=8039931;  
 RA Stevenson B., Bockenstedt L.K., Barthold S.W.;  
 RT "Expression and gene sequence of outer surface protein C of Borrelia  
 burgdorferi reisolated from chronically infected mice."  
 RL Infect. Immun. 62:3568-3571(1994).  
 DR EMBL: U04240; AAC45538.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 SQ SEQUENCE 209 AA; 22393 MW; 3707A47DAA736FCA CRC64;

Query Match 50.1%; Score 901; DB 2; Length 209;  
 Best Local Similarity 98.9%; Pred. No. 4e-35;

Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNSGKGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 60  
 Db 17 ISCNNGKGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 76  
 QY 61 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 120  
 Db 77 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 136  
 Db 121 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180  
 Db 137 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 196  
 QY 181 LTSPV 186  
 Db 197 LTSPV 202

## RESULT 3

09S3P1 PRELIMINARY: PRT: 191 AA.

AC 09S3P1  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OR Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26691;  
 RX MEDLINE=96296448; PubMed=8709845;  
 RA Livey I., Gibbs C.P., Schuster R., Dorner F.;  
 RT "Evidence for lateral transfer and recombination in OSGC variation in

RT Lyme disease Borrelia.";  
 RL Mol. Microbiol. 18:257-269(1995).  
 DR EMBL: I42894; AAB37002.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 191 AA; 20340 MW; 8CA1A64CF17AEDBF CRC64;

Query Match 50.0%; Score 899; DB 2; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 4e-35;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CNNSKRDGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 62  
 Db 1 CNNSKRDGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 60  
 QY 63 KIGNGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 122  
 Db 61 KIGNGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 120  
 QY 123 SEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 182  
 Db 121 SEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180  
 QY 183 SPV 186  
 Db 181 SPV 184

## RESULT 4

031117 PRELIMINARY: PRT: 192 AA.

AC 031117  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSGC.  
 OR Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OC7;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
 RL Submitted (Oct-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF029866; AAB86549.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 192 AA; 20684 MW; 1F0C6BB6E291F6B CRC64;

Query Match 49.5%; Score 889; DB 2; Length 192;  
 Best Local Similarity 97.8%; Pred. No. 1.3e-34;  
 Matches 182; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 60  
 Db 7 ISCNNGKGNASANSADSVKGNLTETSKRTTESNAVVLAVKEVETLLASIDELATKA 66  
 QY 61 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 120  
 Db 67 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAQOCSTEEFTNK 126  
 QY 121 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180  
 Db 127 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 186

OY 181 LTSPV 186  
DB 187 LTSPV 192

## RESULT 5

031122 PRELIMINARY: PRT: 194 AA.

AC 031122:  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
SF STRAIN=OC12;  
Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF029871: AAB86554.1: -  
DR InterPro: IPR001800: Lipoprotein\_6.  
DR Pfam: PF01441: Lipoprotein\_6; 1.  
DR Prodom: PD001149: Lipoprotein\_6; 1.  
FT NON\_TER 1 194  
SQ SEQUENCE 194 AA; 20640 MW; 00A5E6E2D2CE0F7F CRC64;

Query Match 36.9%; Score 663; DB 2; Length 194;  
Best Local Similarity 73.9%; Pred. No. 3.5e-24;

Matches 139; Conservative 22; Mismatches 25; Indels 2; Gaps 2;

OY 1 MACNNSGKGNASANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 60  
DB 7 ISCNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 66  
OY 61 ICKKI-GNNGLEANOSKNTSLSGAVAIISDLIAEKLVNLTKN-EELKEKIDTAKOCSTFT 118  
DB 67 ICKKIQQNGGLAVEAGHNGTLLAGATTISKLITQKLDGKNSKLEKEIKNAKCSDEFT 126  
OY 119 NKLSEHAVLGIDNLTDDNAORAILKHKANKDKGALELEKLFKAVENLSKAAODTLKNAV 178  
DB 127 KLEGGHAGLGIEVNTDENAKKAILITDAKOKGALELEKLFKAVENLSKAAKEMLANSV 186  
OY 179 KELTSPIV 186  
DB 187 KELTSPIV 194

## RESULT 6

P94234 PRELIMINARY: PRT: 193 AA.

AC P94234:  
DT 01-MAY-1997 (TREMblrel. 03, Created)  
DT 01-MAY-1997 (TREMblrel. 03, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
SF STRAIN=28354;  
MEDLINE=96296448; PubMed=8709845;  
RA Livey I., Gibbs C.P., Schuster R., Dornier F.;  
RT "Evidence for lateral transfer and recombination in OSpC variation in Lyme disease Borrelia."  
Mol. Microbiol. 18:257-269(1995).

DR EMBL: I42895; AAB37003.1: -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1 193  
SQ SEQUENCE 193 AA; 20411 MW; 05B68720F061E2A0 CRC64;

Query Match 36.8%; Score 661; DB 2; Length 193;  
Best Local Similarity 74.7%; Pred. No. 4.3e-24;

Matches 139; Conservative 20; Mismatches 25; Indels 2; Gaps 2;

OY 3 CANSKGDNASANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 62  
DB 1 CANSKGDNASANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 60  
OY 63 KKI-GNNGLEANOSKNTSLSGAVAIISDLIAEKLVNLTKN-EELKEKIDTAKOCSTFT 120  
DB 61 KKIQQNGGLAVEAGHNGTLLAGATTISKLITQKLDGKNSKLEKEIKNAKCSDEFT 120  
OY 121 LKSEHAVLGIDNLTDDNAORAILKHKANKDKGALELEKLFKAVENLSKAAODTLKNAV 180  
DB 121 LKSEHAVLGIEVNTDENAKKAILITDAKOKGALELEKLFKAVENLSKAAKEMLANSV 180  
OY 181 LTSPV 186  
DB 181 LTSPV 186

## RESULT 7

Q44705 PRELIMINARY: PRT: 212 AA.

AC Q44705:  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C.  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
SF STRAIN=2591;  
MEDLINE=94041630; PubMed=8225587;  
RA Padua S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;  
RT "Molecular characterization and expression of p23 (OspC) from a North American strain of Borrelia burgdorferi."  
Infect. Immun. 61:5097-5105(1993).  
DR EMBL: U01892; AAA16057.1: -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
SQ SEQUENCE 212 AA; 22270 MW; FB2EF2673A384276 CRC64;

Query Match 36.4%; Score 653.5; DB 2; Length 212;  
Best Local Similarity 72.4%; Pred. No. 1.1e-23;

Matches 139; Conservative 21; Mismatches 29; Indels 3; Gaps 2;

OY 1 MACNNSGKGNASANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 60  
DB 17 ISCNSGKDGNTSANSADSVKGNLTETSKITTESNAVAVLAKVEETLLASIDELATKA 76  
OY 61 ICKKIQQNGGLAVEAGHNGTLLAGATTISKLITQKLDGKNSKLEKEIKNAKCSDEFT 119  
DB 77 ICKKIQQNGGLAVEAGHNGTLLAGATTISKLITQKLDGKNSKLEKEIKNAKCSDEFT 136  
OY 120 LKSEHAVLGIDNLTDDNAORAILKHKANKDKGALELEKLFKAVENLSKAAODTLKNAV 177  
DB 137 KLEGGHAGLGIEVNTDENAKKAILITDAKOKGALELEKLFKAVENLSKAAKEMLANSV 196

OY 178 VKELTSPVHGN 189  
 DB 197 VKELTSPVAEN 208

## RESULT 8

008231 PRELIMINARY; PRT; 202 AA.  
 AC 008231;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OS Borrelia tanukii.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=56146;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RM Masuzawa T., Komikado T., Fukui T., Yanagihara Y.,  
 RT "B. tanukii (Strain F181) ospC, partial cds."  
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AB000354; BAI19087.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 21536 MW; F309F684C68961E7 CRC64;

Query Match 35.9%; Score 646; DB 2; Length 202;  
 Best Local Similarity 70.7%; Pred. No. 2.2e-23;

Matches 133; Conservative 22; Mismatches 31; Indels 2; Gaps 2;

OY 1 MACNNSGKGG-NASANSADSVKGNLTETSKITNESNAVYLAKEVEETLLASIDELATK 59  
 DB 11 ISCNNSGKGDASINPVDSEKSGPLTEISKITDSNAVYLAKEVEETLLASIDELANK 70  
 OY 60 AIGKRIINGNLEANSKNTSLSGAVYISDLIAEKLNVLNK-EELKEKIDTAQOCSTFT 118  
 DB 71 AIGKRIQNNGLDITLSDKNASLSLGAIVISTLETKEKNGKNSDLEKEIEKKKCSAFT 130  
 OY 119 NKLKSEHAVLGDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAADTLKANV 178  
 DB 131 NKLKSHQVLGEAATDDAKKAILKHTYTKDKGAEEFKLFSVEILIKAAOELANSI 190  
 OY 179 KELTSPV 186  
 DB 191 KELTSPV 198

## RESULT 9

044720 PRELIMINARY; PRT; 211 AA.  
 AC 044720;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE OSPC.  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RM STRAIN-25015;  
 RC MEDLINE-95154673; Pubmed-7851744;  
 RA Stevenson B., Barthold S.W.;  
 RT "Expression and sequence of outer surface protein C among North  
 American isolates of Borrelia burgdorferi."  
 RL FEMS Microbiol. Lett. 124:367-372(1994).  
 DR EMBL: U04282; AAC45540.1; -

DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 SQ SEQUENCE 211 AA; 22603 MW; 63984BA6D8743ED5 CRC64;

Query Match 35.9%; Score 646; DB 2; Length 211;  
 Best Local Similarity 72.3%; Pred. No. 2.3e-23;  
 Matches 136; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

OY 1 MACNNSGKDG-NASANSADSVKGNLTETSKITNESNAVYLAKEVEETLLASIDELATK 59  
 DB 17 ISCNNSGKGDASINPVDSEKSGPLTEISKITDSNAVYLAKEVEETLLASIDELATK 76  
 OY 60 AIGKRI-GNNGLEANSKNTSLSGAVYISDLIAEKLNVLNK-EELKEKIDTAQOCSTFT 118  
 DB 77 AIGKRIHQNNGLDTEENNNGSLGAVYISLTQKLGLENEELKEKIAVKKCSSEFT 136  
 OY 119 NKLKSEHAVLGDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAADTLKANV 178  
 DB 137 NKLKSHTELGKQADDDAKKAILKHTYTKDKGAEEIDKLFKAVENLSKAKEMLNSV 196  
 OY 179 KELTSPV 186  
 DB 197 KELTSPV 204

## RESULT 10

09R08 PRELIMINARY; PRT; 201 AA.  
 AC 09R08;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_Taxid=139;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RM STRAIN-386';  
 RC MEDLINE-20002545; Pubmed-10531219;  
 RA Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;  
 RT "Population dynamics of a naturally occurring heterogeneous mixture of  
 Borrelia burgdorferi clones."  
 RL Infect. Immun. 67:5709-5716(1999).  
 DR EMBL: AF074465; AAD23912.1;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 201  
 SQ SEQUENCE 201 AA; 21457 MW; 806F198295101B07 CRC64;

Query Match 35.8%; Score 644; DB 2; Length 201;  
 Best Local Similarity 73.4%; Pred. No. 2.7e-23;  
 Matches 135; Conservative 22; Mismatches 25; Indels 2; Gaps 2;

OY 1 MACNNSGKDG-NASANSADSVKGNLTETSKITNESNAVYLAKEVEETLLASIDELATK 60  
 DB 17 ISCNNSGKGDASINPVDSEKSGPLTEISKITDSNAVYLAKEVEETLLASIDELATK 76  
 OY 61 IGKRI-GNNGLEANSKNTSLSGAVYISDLIAEKLNVLNK-EELKEKIDTAQOCSTFT 118  
 DB 77 IGKRIQNNGLDTEENNNGSLGAVYISLTQKLGLENEELKEKIAVKKCSSEFT 136  
 OY 119 NKLKSEHAVLGDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAADTLKANV 178  
 DB 137 KKLGEHAQVLGENTVDENAKKAILITTDAAKDKGAEEIDKLFKAVENLSKAKEMLANSV 196

OY 179 KELL 182  
DB 197 KELL 200

## RESULT 11

P94237 PRELIMINARY: PRT: 193 AA.

AC 01-MAY-1997 (TREMELREL. 03, Created)  
DT 01-MAY-1997 (TREMELREL. 03, last sequence update)  
DE 01-JUN-2001 (TREMELREL. 17, last annotation update)  
GN OUTR SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_Taxid=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=25015;  
PX MEDLINE=96296448; PubMed=8709845;  
RT Livey I., Gibbs C.P., Schuster R., Dorner F.;  
"Evidence for lateral transfer and recombination in OspC variation in  
Lyme disease Borrelia."  
RL MOL. Microbiol. 18:257-269(1995).  
DR EMBL: L42898; AAB37007.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT SEQUENCE 193 193 MW: C9500959E13590D CRC64;

Query Match 35.7%; Score 642; DB 2; Length 193;  
Best Local Similarity 72.6%; Pred. No. 3.2e-23;  
Matches 135; Conservative 16; Mismatches 33; Indels 2; Gaps 2;

OY 3 CUNSGKDG-N-ASANSADSVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 61  
DB 1 CUNSGKDGNAASNPADSVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 60  
OY 62 GKRI-GNNGLEANOSKNTSILSGAVAIISDLIAEKLAVLKNEELKEKIDTAOKCSTFTNK 120  
DB 61 GKRIHONNGIDETENHNGSLAGAVAIISLTITOKLGLKNEELKEKIAAVKCSSEFTNK 120  
OY 121 LKSEHAVLGLDNLTPDNAORAILKHKANKGAAELEKLFKAVENTLSKAQDTLKNAYKE 180  
DB 121 LKSSHTELCKQADODDADAKKAILRTINTKDKGAELEKLFKPYENLSKAKEMLSNVKE 180  
OY 181 LTPSPV 186  
DB 181 LTPSPV 186  
RESULT 12  
OY 044726 PRELIMINARY: PRT: 191 AA.  
AC 044726;  
DT 01-NOV-1996 (TREMELREL. 01, Created)  
DT 01-NOV-1996 (TREMELREL. 01, last sequence update)  
DE 01-JUN-2001 (TREMELREL. 17, last annotation update)  
GN OUTR SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_Taxid=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=297;  
PX MEDLINE=94314437; PubMed=8039891;  
RT Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;  
"Humoral immune response to outer surface protein C of Borrelia  
burgdorferi in Lyme disease: role of the immunoglobulin M response in

RT the serodiagnosis of early infection.";

RL Infect. Immun. 62:3213-3221(1994).

DR EMBL: U08284; AAA21460.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT SEQUENCE 191 191 MW: C49A4030F0A28717 CRC64;

Query Match 35.6%; Score 639; DB 2; Length 191;  
Best Local Similarity 73.9%; Pred. No. 4.4e-23;  
Matches 136; Conservative 20; Mismatches 26; Indels 2; Gaps 2;

OY 5 NSGKDGNAASNSADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 64  
DB 1 NSGKDGNAASNSADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 60  
OY 65 I-GNNGLEANOSKNTSILSGAVAIISDLIAEKLAVLK-N-EELKEKIDTAOKCSTFTNK 122  
DB 61 IONGGLAVAGAHNGTLLAGAVITISKITOKLGLKNEELKEKIAAVKCSSEFTNK 120  
OY 123 SEHAVLGLDNLTPDNAORAILKHKANKGAAELEKLFKAVENTLSKAQDTLKNAYKE 182  
DB 121 GEHAGLGIENVTDENAKKAILITDAADKGAELEKLFKAVENTLSKAQDTLKNAYKE 180  
OY 183 SPV 186  
DB 181 SPV 184

RESULT 13  
OY 09R7B1 PRELIMINARY: PRT: 192 AA.  
AC 09R7B1;  
DT 01-MAY-2000 (TREMELREL. 13, Created)  
DT 01-MAY-2000 (TREMELREL. 13, last sequence update)  
DE 01-JUN-2001 (TREMELREL. 17, last annotation update)  
GN OUTR SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_Taxid=139;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=TESTS;  
PX MEDLINE=97478003; PubMed=9336916;  
RT Ras N.M., Postic D., Foretz M., Baranton G.;  
"Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
U.S.A.'?";  
Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=TESTS;  
PX Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U91798; AAB81895.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT SEQUENCE 192 192 MW: 6770502A20AAAF64 CRC64;

Query Match 34.7%; Score 623.5; DB 2; Length 192;  
Best Local Similarity 69.5%; Pred. No. 2.3e-22;  
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

OY 1 MACNSGKDGNAASNSADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 60  
DB 3 ISGKNSKDKDNTSNAASADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 62  
OY 61 IGKRI-GNNGLEANOSKNTSILSGAVAIISDLIAEKLAVLKNEELKEKIDTAOKCSTFTNK 119

```

Db 63 IGGKHQNNGLDTEENNNGSLAGAVAIISLILKQKLDGKNEGLKEDIKIDAKKCSFTFN 122
QY 120 KLKSEHAVLGLDNLTDNAORAILKKHANKDGAELKLFKAVENLSKAADTLKNAVY 179
Db 123 KLEKEHTDLGKEGVTDADAKAELIKTNGTKGAELIKLFESVEYLSRAAKEMLANSVK 182
QY 180 ELTSPVIV 186
Db 183 ELTSPVIV 189

RESULT 14
Q9RR53 PRELIMINARY; PRT; 193 AA.
AC Q9RR53;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC1;
RA MEDLINE=99091544; PubMed=9872945;
Wang J.N., Dykhuizen D.E., Olu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of oSCP in a local population of Borrelia
burgdorferi sensu stricto."
RL Genetics 151:15-30(1999).
DR EMBL; AF029860; AAB86543.1;
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20502 MW; 5EFD5AF8986D1E CRC64;

```

```

Query Match 34.7%; Score 623.5; DB 2; Length 193;
Best Local Similarity 69.5%; Pred. No. 2.3e-22;
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

```

```

QY 1 MACNNSGKDGNAANSADSVKGPMLTEISKRTESNAVLAKEVEFTLASIDELATKA 60
7 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEFTLASIDELATKA 66
QY 61 IGKRI-GNNGLEANOSKNTSLSGAVAISDLIAEKLNVLAKEVEFTLASIDELATKA 119
Db 67 IGGKHQNNGLDTEENNNGSLAGAVAIISLILKQKLDGKNEGLKEDIKIDAKKCSFTFN 126
QY 120 KLKSEHAVLGLDNLTDNAORAILKKHANKDGAELKLFKAVENLSKAADTLKNAVY 179
Db 127 KLEKEHTDLGKEGVTDADAKAELIKTNGTKGAELIKLFESVEYLSRAAKEMLANSVK 186
QY 180 ELTSPVIV 186
Db 187 ELTSPVIV 189

RESULT 15
Q31123 PRELIMINARY; PRT; 185 AA.
AC Q31123;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC13;
RA Wang J.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RT Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF029872; AAB86555.1;
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

```

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC13;
RA Wang J.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RT Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF029872; AAB86555.1;
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

```

```

Query Match 34.7%; Score 623; DB 2; Length 185;
Best Local Similarity 72.1%; Pred. No. 2.3e-22;
Matches 129; Conservative 24; Mismatches 24; Indels 2; Gaps 2;

```

```

Search completed: March 18, 2002, 10:10:55
Job time: 978 sec

```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:33 ; Search time 118.14 Seconds  
(without alignments)  
230.734 Million cell updates/sec

Title: US-09-596-746A-42  
1798

Perfect score: 1 MACNNSCKDGNASNADES.....AVENLAKAKEMILANSVKEL 368

Sequence: 1 MACNNSCKDGNASNADES.....AVENLAKAKEMILANSVKEL 368

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Maximum number of hits satisfying chosen parameters: 522463

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

A.Geneseq\_1101:\*

- 1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1793	99.7	22	AAB62719
2	1790	99.6	22	AAB62735
3	1554.5	86.6	22	AAB62718
4	1554.5	86.6	22	AAB62734
5	1538	85.5	22	AAB62716
6	1530	85.1	22	AAB62736
7	1485.5	82.6	22	AAB62713
8	1484	82.5	22	AAB62711
9	1481	82.4	22	AAB62729
10	1480	82.3	22	AAB62716
11	1479.5	82.3	22	AAB62739

Result No.	Score	Query Match length	ID	Description
12	1477	82.1	393	22 AAB62731
13	1250	69.5	378	22 AAB62712
14	1248.5	69.4	374	22 AAB62710
15	1245.5	69.3	398	22 AAB62728
16	1244.5	69.2	370	22 AAB62715
17	1244	69.2	401	22 AAB62738
18	1241.5	69.0	394	22 AAB62730
19	1230	67.9	369	22 AAB62714
20	1217	67.7	393	22 AAB62732
21	1167.5	64.9	560	22 AAB62724
22	1159.5	64.5	384	22 AAB62726
23	1156.5	64.3	408	22 AAB62737
24	1144	63.6	378	22 AAB62725
25	1136	63.2	401	22 AAB62733
26	1101.5	61.3	410	22 AAB62740
27	1100.5	61.2	386	22 AAB62727
28	901	50.1	191	22 AAB62705
29	899	50.0	184	22 AAB62702
30	889	50.0	191	15 AAB60884
31	889	49.4	191	22 AAB62706
32	876	48.7	193	22 AAB62709
33	699	38.9	587	16 AAR75746
34	698.5	38.8	466	16 AAR75739
35	665	37.0	466	16 AAR75740
36	657.5	36.6	212	16 AAM11934
37	651.5	36.2	194	15 AAR60885
38	644.5	35.8	192	15 AAB60886
39	644.5	35.8	192	22 AAB62703
40	644.5	35.8	210	16 AAM11935
41	640.5	35.6	190	22 AAB62707
42	640	35.6	193	15 AAR60888
43	640	35.6	211	18 AAM41822
44	639.5	35.6	192	15 AAR60889
45	639.5	35.6	210	16 AAR75727

## ALIGNMENTS

RESULT 1

AB62719 standard; Protein; 367 AA.

AC AAB62719;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospC protein SEQ ID NO: 42.

XX Borrelia: ospC; Lyme disease; vaccine; chimeric protein; tick.

XX Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000MO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNV ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuzen D, Luft BJ, Gomes-Solecki M;

XX WPI: 2001-050113/06.

XX N-PSDB: AAR29023.

XX Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

XX Claim 43: Page 102-103; 160pp; English.

PS The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 367 AA:

Query Match 99.7%; Score 1793; DB 22; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 1,2e-108;  
 Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 ACNNGSGKDGMSANSADSVKGPNTLTKSKITSSNAVLAKEVEETLLASIDELATKAI 61
   1 acnngsgkdgmsansadesvkgpnlteiskkltesnavlavkevelasidelatka1 60
QY 62 GKRIQNGCLEANQSKNTSLSGAVAIISDLIAEKLVNFKNEELKEKIDTAQOCSTEFNTKL 121
   61 gkriqngcleangskntslsgayaisdliaeklnvfkneelkekidtaqocstefntkl 120
QY 122 KSEHAVIGLDLITDDNQRAILKHKANKDGALEKLFRAVENLSKRAODTLKNAVKEL 181
   121 ksehavigldlittddnqraillkhhankdgaalekllfavenlskaagdtlknavekl 180
QY 182 TSPIVHGNSGDKGNTSANSADSVKGPNTLTKSKITTESNAVLAKEVEETLLASIDEL 241
   181 tspivhgnsdkgntsansadesvkgpnlteiskkltesnavlavkeletllasidel 240
QY 242 ATTAIGKKIOQNGLAIVEAGHNGTLAAGATTISKLITQKIDGKNSKLEKIEKNAKKS 301
   241 attaigkkiqngglaveaghnngtllagattysklitqkldgknsklekienakks 300
QY 302 EDFTKKLEGEHAGLGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAAKEML 361
   301 edftkklegelhaglgienvtdenakkailitdaakdkgaalekllfavenlakaakeml 360
QY 362 ANSVKEL 368
   361 ansvkel 367
Db

```

RESULT 2

AA62735 standard: Protein; 391 AA.

AC AAB62735;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 74.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNY ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dattwyler RJ, Selinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI: 2001-050113/06.

DR N-PDB; AAF29039.

XX

PT Compositions of ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

XX to Lyme disease -

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX (UINY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
PI  
XX WPI: 2001-050113/06.  
DR N-PSDB; AAF29022.  
XX  
XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -  
XX  
XX  
PS Claim 43; Page 99-100; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

Sequence 368 AA;

Query Match 86.6%; Score 1557.5; DB 22; Length 368;  
Best Local Similarity 87.5%; Pred. No. 2.1e-93;  
Matches 321; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 2 ACNNSGKDNASANSADSVKGNLTETISKITTESNAVVLAVKEVETLLASIDELATKAI 61  
DB 1 acnnskgdnasansadesvkgpnltetiskitesnavvlavkevetllasidelatka 60  
OY 62 GKRIKGNGLANOSKNTSLSGAVAISDLIAEKLNVLNKELEKIDTAKOCSTETFNKL 121  
DB 61 gkrikgngleangskntsllsgavalsdliiaeKlnvlnkeelkeIdtakqstefnkl 120  
OY 122 KSEHNAVIGLNDITDDNAQRAILKKHANKDGALEKLFRAVENLSKAADDTLKNAYKEL 181  
DB 121 ksehnavlgldntddnaqraillkhanKdkgaaleklfravenlskaadqtlknaykel 180  
OY 182 TSPIVHNGNSGKDNNTSANSADSVKGNLTETISKITTESNAVVLAVKEVETLLASIDEL 241  
DB 181 tspivhngnsgkdngntsansadesvkgpnltetiskitesnavvlavkevetllasidel 240  
OY 242 AFRAIGKKTIOQNGLAVEAGHNGTLLAGAYTISKLTOKLDGKNSKLEKIEKNARKCS 301  
DB 241 a-fraigkktioqnglaveaghnGtllagaytiskltokldgknskLeKieknarkcs 299  
OY 302 EDFTKKLEGEHAOLGIENTYDENAKKAILITDAKDKGALEKLFRAVENLAKAAKEML 361  
DB 300 eedfaklkgehtldlgkyevtdnnaKkailtkndcktgadeleklfesvknlskaakeml 359  
OY 362 ANSVKEL 368  
DB 360 tnsvkel 366

RESULT 4  
AAB62734  
ID AAB62734 standard; Protein; 392 AA.  
XX  
XX AAB62734;  
XX  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospc protein SEQ ID NO: 72.  
XX  
XX  
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
XX Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX  
XX WO200078966-A1.

XX 28-DEC-2000.  
XX  
XX 19-JUN-2000; 2000WO-US16915.  
XX  
XX 18-JUN-1999; 99US-0140042.  
XX  
XX (UINY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
PI  
XX WPI: 2001-050113/06.  
DR N-PSDB; AAF29038.  
XX  
XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -  
XX  
XX  
PS Claim 43; Page 139-140; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

Sequence 392 AA;

Query Match 86.5%; Score 1554.5; DB 22; Length 392;  
Best Local Similarity 87.2%; Pred. No. 3.5e-93;  
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

OY 2 ACNNSGKDNASANSADSVKGNLTETISKITTESNAVVLAVKEVETLLASIDELATKAI 61  
DB 25 acnnskgdnasansadesvkgpnltetiskitesnavvlavkevetllasidelatka 84  
OY 62 GKRIKGNGLANOSKNTSLSGAVAISDLIAEKLNVLNKELEKIDTAKOCSTETFNKL 121  
DB 85 gkrikgngleangskntsllsgavalsdliiaeKlnvlnkeelkeIdtakqstefnkl 144  
OY 122 KSEHNAVIGLNDITDDNAQRAILKKHANKDGALEKLFRAVENLSKAADDTLKNAYKEL 181  
DB 121 ksehnavlgldntddnaqraillkhanKdkgaaleklfravenlskaadqtlknaykel 180  
OY 182 TSPIVHNGNSGKDNNTSANSADSVKGNLTETISKITTESNAVVLAVKEVETLLASIDEL 241  
DB 181 tspivhngnsgkdngntsansadesvkgpnltetiskitesnavvlavkevetllasidel 240  
OY 242 AFRAIGKKTIOQNGLAVEAGHNGTLLAGAYTISKLTOKLDGKNSKLEKIEKNARKCS 301  
DB 241 a-fraigkktioqnglaveaghnGtllagaytiskltokldgknskLeKieknarkcs 299  
OY 302 EDFTKKLEGEHAOLGIENTYDENAKKAILITDAKDKGALEKLFRAVENLAKAAKEML 361  
DB 300 eedfaklkgehtldlgkyevtdnnaKkailtkndcktgadeleklfesvknlskaakeml 359  
OY 362 ANSVKEL 368  
DB 384 tnsvkel 390

RESULT 5  
AAB62717  
ID AAB62717 standard; Protein; 368 AA.  
XX  
XX AAB62717;  
XX  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospc protein SEQ ID NO: 38.  
XX

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
XX Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX WO200078966-A1.  
XX 28-DEC-2000.  
XX 19-JUN-2000; 2000MO-US16915.  
XX 18-JUN-1999; 99US-0140042.  
XX (UANY ) UNIV NEW YORK STATE RES FOUND.  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX MPI; 2001-050113/06.  
XX N-PSDB; AAF29021.  
XX  
XX Compositions of ospC polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -  
XX  
XX Claim 43; Page 97; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospC proteins and  
XX chimeric ospC proteins from members of the Borrelia genus. These may be  
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
XX vaccines against Borrelia infection, which is spread by ticks and leads  
XX to Lyme disease.  
XX  
SQ Sequence 368 AA:  
  
Query Match 85.5%; Score 1538; DB 22; Length 368;  
Best Local Similarity 86.7%; Pred. No. 3.8e-92;  
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;  
  
QY 1 MACNSGKDGMSANSADSVKGPMLTEISKRTTESNAVAVLAVKEVETLLASIDELATKA 60  
DB 1 macnsgkdgnasansadesvkgpnlteiskrttesnavlavkevettllasidelatka 60  
QY 61 ICKKIGNGLEANSKNTSLSGAYASDLIAEKLNVLKNEELKEKIDTAKOCSTEFNTK 120  
DB 61 ICKKIGNGLEANSKNTSLSGAYASDLIAEKLNVLKNEELKEKIDTAKOCSTEFNTK 120  
QY 121 LKSEHAVIGLDNLTDNAORAILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180  
DB 121 lksenhaviglndltddnagrailkhhankdkgaaleklfkavenlskaagdtlknavek 180  
QY 181 LTPSPVHGNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVAVLAVKEVETLLASIDE 240  
DB 181 ltpspvhgnsrkdgnasnsadesvkgpnlteiskrttesnavlavkevettllasidel 240  
QY 241 LATKAIIGKRIQONGGLAVEAGHNGTLGAYTISKLTIOKLDGLKNSKLEKTEENAKKC 300  
DB 241 latkaiigkriqongglaveaghnngtllagaytisikltioqldglknskleteenakkc 300  
QY 301 SEDFTKKEGHEHAQIGIENVTDENAKKAILITDAKDKGAELKLFKAVENLSKAAKEM 360  
DB 301 sedftkkeghehaqigienvtidenakkailittdakdkgaaleklfkavenlskaagdt 360  
QY 361 LANSYKEL 368  
DB 361 lnavkel 366  
RESULT 6  
ID AAB62736 standard; Protein; 391 AA.  
XX

AC AAB62736;  
XX 03-APR-2001 (first entry)  
XX  
XX Borrelia sp chimeric ospC protein SEQ ID NO: 76.  
XX  
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
XX Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX WO200078966-A1.  
XX 28-DEC-2000.  
XX 19-JUN-2000; 2000MO-US16915.  
XX 18-JUN-1999; 99US-0140042.  
XX (UANY ) UNIV NEW YORK STATE RES FOUND.  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
XX MPI; 2001-050113/06.  
XX N-PSDB; AAF29040.  
XX  
XX Compositions of ospC polypeptides from strains of Borrelia which cause  
XX Lyme disease are used to immunize animals and detect immune responses  
XX to Lyme disease -  
XX  
XX Claim 43; Page 144-145; 160pp; English.  
XX  
XX The present invention provides compositions comprising ospC proteins and  
XX chimeric ospC proteins from members of the Borrelia genus. These may be  
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
XX vaccines against Borrelia infection, which is spread by ticks and leads  
XX to Lyme disease.  
XX  
SQ Sequence 391 AA:  
  
Query Match 85.1%; Score 1530; DB 22; Length 391;  
Best Local Similarity 86.4%; Pred. No. 1.4e-91;  
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;  
  
QY 2 ACNNSGKDGMSANSADSVKGPMLTEISKRTTESNAVAVLAVKEVETLLASIDELATKA 61  
DB 2 acnnsdgkdnasansadesvkgpnlteiskrttesnavlavkevettllasidelatka 61  
QY 62 GKKTIGNGLEANSKNTSLSGAYASDLIAEKLNVLKNEELKEKIDTAKOCSTEFNTK 121  
DB 62 gkktigngleangskntslsgayaisdliaeklnvlnkeelkekiddakqstefntkl 121  
QY 122 KSEHAVIGLDNLTDNAORAILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 181  
DB 122 ksenhaviglndltddnagrailkhhankdkgaaleklfkavenlskaagdtlknavek 181  
QY 182 TSPVHGNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVAVLAVKEVETLLASIDE 241  
DB 182 tspvhgnsrkdgnasnsadesvkgpnlteiskrttesnavlavkevettllasidel 241  
QY 242 ATKAIIGKRIQONGGLAVEAGHNGTLGAYTISKLTIOKLDGLKNSKLEKTEENAKKC 301  
DB 242 atkaiigkriqongglaveaghnngtllagaytisikltioqldglknskleteenakkc 301  
QY 302 EDFTKKEGHEHAQIGIENVTDENAKKAILITDAKDKGAELKLFKAVENLSKAAKEM 361  
DB 302 edftkkeghehaqigienvtidenakkailittdakdkgaaleklfkavenlskaagdt 361  
QY 362 ANSYKEL 368  
DB 362 knavkel 389

## RESULT 7

AAB62713  
ID AAB62713 standard; Protein: 377 AA.

AC AAB62713;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospC protein SEQ ID NO: 30.

KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO20078966-A1.

PD 28-DEC-2000.

PR 19-JUN-2000; 2000WO-US16915.

PA (UNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Sehnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR MPI: 2001-050113/06.

XX N-PSDB; AAF29017.

PT Compositions of ospC polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 86-87; 160pp; English.

CC The present invention provides compositions comprising ospC proteins and  
CC chimeric ospC proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 377 AA;

Query Match 82.6%; Score 1485.5; DB 22; Length 377;

Best Local Similarity 82.5%; Pred. No. 9.8e-89; Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

QY 1 MACNNSGKDGNSANSADSVKGNLTETSKITTESNAVLAVKEVETLLASIDELATKA 60  
DB 1 macnsgkdgntsanadesvkgpnlteinkktdsnavlavkeveallsidelaaka 60  
QY 61 ICKKI-GNNGLEANOSKNTSLSGAIVASDLIAEKLVKNEELKEKIDTAKOCSTEFN 119  
DB 61 ICKKI-GNNGLEANOSKNTSLSGAIVASDLIAEKLVKNEELKEKIDTAKOCSTEFN 119  
QY 120 KKESEHAVLGLDNLTDNORAILKKHANKDGAELKLFKAVENLSKAADOTLKNAV 129  
DB 120 KKESEHAVLGLDNLTDNORAILKKHANKDGAELKLFKAVENLSKAADOTLKNAV 129  
QY 121 KKEHKLIDYGEVTDADAKEAIIKNGCTKGAELGLFSEVYLSKAKEMLANSVK 180  
DB 121 KKEHKLIDYGEVTDADAKEAIIKNGCTKGAELGLFSEVYLSKAKEMLANSVK 180  
QY 180 ELTSPIVHG-----NNSGKDGNTSANSADSVKGNLTETSKITTESNAVLAVKEI 231  
DB 180 ELTSPIVHG-----NNSGKDGNTSANSADSVKGNLTETSKITTESNAVLAVKEI 231  
QY 232 ETLIASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGATTISKLTTQKIDGLKNSKK 291  
DB 232 ETLIASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGATTISKLTTQKIDGLKNSKK 291  
QY 241 ETLIASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGATTISKLTTQKIDGLKNSKK 300  
DB 241 ETLIASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGATTISKLTTQKIDGLKNSKK 300  
QY 292 EKLENKKSEDFTKKLEGEHAGLGIENTVDENAKKAILITDAKKGAELKLFKAVE 351  
DB 292 EKLENKKSEDFTKKLEGEHAGLGIENTVDENAKKAILITDAKKGAELKLFKAVE 351

DB 301 EKENAKKSEDFTKKLEGEHAGLGIENTVDENAKKAILITDAKKGAELKLFKAVE 360

QY 352 NLAKAKKEMIANSVKEL 368

DB 361 nlakakkemiansvkel 377

## RESULT 8

AAB62711  
ID AAB62711 standard; Protein: 373 AA.

AC AAB62711;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospC protein SEQ ID NO: 26.

KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO20078966-A1.

PD 28-DEC-2000.

PR 19-JUN-2000; 2000WO-US16915.

PA (UNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Sehnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR MPI: 2001-050113/06.

XX N-PSDB; AAF29015.

CC Compositions of ospC polypeptides from strains of Borrelia which cause  
CC Lyme disease are used to immunize animals and detect immune responses  
CC to Lyme disease -

PS Claim 43; Page 81; 160pp; English.

CC The present invention provides compositions comprising ospC proteins and  
CC chimeric ospC proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

XX Sequence 373 AA;

Query Match 83.4%; Score 1484; DB 22; Length 373;

Best Local Similarity 83.4%; Pred. No. 1.2e-88; Matches 311; Conservative 22; Mismatches 34; Indels 6; Gaps 2;

QY 2 ACNNSGKDGNSANSADSVKGNLTETSKITTESNAVLAVKEVETLLASIDELATKA 61  
DB 2 acnsgkdgntsanadesvkgpnlteiskktdsnavlavkeveallsidelaaka 60  
QY 62 GKKI-GNNGLEANOSKNTSLSGAIVASDLIAEKLVKNEELKEKIDTAKOCSTEFN 120  
DB 62 GKKI-GNNGLEANOSKNTSLSGAIVASDLIAEKLVKNEELKEKIDTAKOCSTEFN 120  
QY 121 KKESEHAVLGLDNLTDNORAILKKHANKDGAELKLFKAVENLSKAADOTLKNAV 180  
DB 121 KKESEHAVLGLDNLTDNORAILKKHANKDGAELKLFKAVENLSKAADOTLKNAV 180  
QY 181 LTPSPIVHG-----NNSGKDGNTSANSADSVKGNLTETSKITTESNAVLAVKEITLL 235  
DB 181 LTPSPIVHG-----NNSGKDGNTSANSADSVKGNLTETSKITTESNAVLAVKEITLL 235

QY 236 ASIDELATKAIGKRIQONGGLAVEAGHNGTLLAGAVYTSKLTITOKLDGLKSEKLEKIE 295  
 DB 241 asidelatkaigkriqngglaveaghgntlagaYtskltitqkldglkseklikkie 300  
 QY 296 NAKKSEDFTKLGEHAQOLCIENVTDENAKKAILITDAKDGALEKLFKAVENIAK 355  
 DB 301 nakksedftkllegehaqgienvtdenakkailitdaakdgaaelekikfaveniak 360  
 QY 356 AAKEMIANSVKEL 368  
 DB 361 aakemiansvkel 373

RESULT 9  
 AAB62729  
 ID AAB62729 standard; Protein: 397 AA.

QY AAB62729;  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SFQ ID NO: 62.  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 PD 28-DEC-2000.  
 PF 19-JUN-2000; 2000WO-US16915.  
 PR 18-JUN-1999; 99US-0140042.  
 PA (UYNV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PSDB; AAF29033.  
 XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PI to Lyme disease -  
 PS Claim 43; Page 125-126; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 397 AA;

QY 2 ACNNSGKDGANSANSADESVKGPULTEISKITTESNAVLAKEVEETLLASIDELATKAI 61  
 DB 25 scnmsgdgntsansadesvkgpultelskittsdnsavllavkeveallssideiaakai 84  
 QY 62 GKRI-GNNGLEANOSKMTSLISGAYAISDLIAEKLNVKNELEKIDTAKOCSTEEFNK 120  
 DB 85 gkrihgngldeteynhngslilagayaisdlkqldgkneglekidaakccsetlntk 144  
 QY 121 LKSHAVALGIDNLTDNNAQRALIKKHANKDKGALEKLFKAVENLSKAQDTLKNVKE 180

DB 145 lkehtldygeyvdadakeaalktngltkgaeeigrlfsvvevlskaakemiansvke 204  
 QY 181 LTSPILVHS-----NNSGKDGTSANSANSADESVKGPULTEISKITTESNAVLAKEIEITLL 235  
 DB 205 ltspsvaespmvnsygdntsansadesvkgpultelskittsdnsavllavkeieitll 264  
 QY 236 ASIDELATKAIGKRIQONGGLAVEAGHNGTLLAGAVYTSKLTITOKLDGLKSEKLEKIE 295  
 DB 265 asidelatkaigkriqngglaveaghgntlagaYtskltitqkldglkseklikkie 324  
 QY 296 NAKKSEDFTKLGEHAQOLCIENVTDENAKKAILITDAKDGALEKLFKAVENIAK 355  
 DB 325 nakksedftkllegehaqgienvtdenakkailitdaakdgaaelekikfaveniak 384  
 QY 356 AAKEMIANSVKEL 368  
 DB 385 aakemiansvkel 397

RESULT 10  
 AAB62716  
 ID AAB62716 standard; Protein: 369 AA.

QY AAB62716;  
 DT 03-APR-2001 (first entry)  
 DE Borrelia sp chimeric ospc protein SFQ ID NO: 36.  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 PN WO200078966-A1.  
 PD 28-DEC-2000.  
 PF 19-JUN-2000; 2000WO-US16915.  
 PR 18-JUN-1999; 99US-0140042.  
 PA (UYNV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PSDB; AAF29020.  
 XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PI to Lyme disease -  
 PS Claim 43; Page 94-95; 160pp; English.  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 369 AA;

QY 2 ACNNSGKDGANSANSADESVKGPULTEISKITTESNAVLAKEVEETLLASIDELATKAI 61  
 DB 1 acnmsgdgntsansadesvkgpultelskittsdnsavllavkeveallssideiaakai 59



SO Sequence 400 AA;

82.38; Score 1479.5; DB 22; Length 400;

385 lakaakemlansvkel 400

XX  
XX  
El

CC vaccines against *Borrelia* infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 393 AA;

Query Match 82.1%; Score 1477; DB 22; Length 393;  
 Best Local Similarity 83.5%; Pred. No. 3.7e-88;  
 Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

OY 2 ACNNSGKDGNSANSADSVKGNLTETISKRTESNAVLAWEVETLLASIDELATKAI 61  
 Db 25 scnskdkdntnsadsesvkgnltetiskrtesnaavlaveveallssidela-kai 83  
 OY 62 GKRTGNGG-LEANQSKNTSLSGAVATSDLAETKLVLANEE-LKEKIDTAKQCTEFTN 119  
 Db 84 gkrtkngdsidneanrnesllagaystlctqklskngseglkekaakcseft 143  
 OY 120 KIKSEHAVLGIDNLTDDNQRATLKH-ANKDKGAELKELKAVENLSKAADTLKANAV 178  
 Db 144 kldmhaqldgvtidenakkaikanaagkdyeeleksgslskskaakemlansv 203  
 OY 179 KELTSPVHGNNSGKDGNTSANSADSVKGNLTETISKRTESNAVLAWEVETLLASI 238  
 Db 204 keltspvvhgnmsgkdgntsansadesvkgpnlteiskrtesnavlavkeltellasi 263  
 OY 239 DELATATIKKTIQONGCLAVEAGHNGCTLLAGAVTTSKLTORLIDGLKNSKLEKIEENAK 298  
 Db 264 delatatkiktkiqngglaveagngctllagaytiskltqldglknskilekienak 323  
 OY 299 KCSSEDTKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKELKAVENLSKA 358  
 Db 324 kcsedtkklegehaqlgientvdenakkailltdaakdgaaelekfkavenlskaak 383  
 OY 359 EMLANSVKEL 368  
 Db 384 emlansvkel 393

## RESULT 13

AAB62712  
 ID AAB62712 standard; Protein: 378 AA.

AC AAB62712;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 28.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

XX N-PSDB; AAF29016.

Compositions of OspC polypeptides from strains of Borrelia which cause  
 Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -

PS Claim 43; Page 83-84; 160pp; English.

CC The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 378 AA;

Query Match 69.5%; Score 1250; DB 22; Length 378;  
 Best Local Similarity 70.3%; Pred. No. 1.7e-73;  
 Matches 265; Conservative 43; Mismatches 59; Indels 10; Gaps 3;

OY 1 MACNNSGKDGNSANSADSVKGNLTETISKRTESNAVLAWEVETLLASIDELATKA 60  
 Db 1 macnnskgdgnlsansadesvkgpnlteinkrtesnavllaveveallssidelaaka 60  
 OY 61 IGKRT-GNNGLEANQSKNTSLSGAVATSDLAETKLVLANEELKEDTAKQCTEFTN 119  
 Db 61 igkrtghngldtenhngslilagayaistlkqldgkngelkedaakcseftn 120  
 OY 120 KIKSEHAVLGIDNLTDDNQAQRAILKHKANKDKGAELKELKAVENLSKAADTLKANAV 179  
 Db 121 kkehtldlgkyvtdadakaalkangtktkgaeelgklesvevlskaakemlansvk 180  
 OY 180 ELTSPVHNG-----NNSGKDGNTSANSADSVKGNLTETISKRTESNAVLAWEI 231  
 Db 181 eltspvvaeppkpsmynsgkdgntsansadesvkgpnlteiskrtesnavlavkev 240  
 OY 232 ETLASIDELATKAIIGKRTIQONGCLAVEAGHNGCTLLAGAVTTSKLTORLIDGLKNSK 291  
 Db 241 etlltsidela-kaiqkikndvsldeadhngslisgaylslntkksaisaldsgelk 299  
 OY 292 EKIEANAKCSEDTKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKELKAVE 351  
 Db 300 aeieakkkcseefaklgehtldlgkyvtdadnakkaikntndktkgadelkelfesvk 359  
 OY 352 NLAKRAKEMLANSVKEL 368  
 Db 360 nlskaakemlansvkel 376

## RESULT 14

AAB62710  
 ID AAB62710 standard; Protein: 374 AA.

AC AAB62710;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 24.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY ) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

DR N-PSDB: AAF29014.

XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -

PS Claim 43; Page 78-79; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 374 AA;

Query Match 69.4%; Score 1248.5; DB 22; Length 374;  
 Best Local Similarity 71.0%; Pred. No. 2,1e-73;  
 Matches 265; Conservative 42; Mismatches 59; Indels 7; Gaps 3;

DB 2 ACNNSGKDGNSANSADSVKGNPLTEISKTTESNAVVLAVKEVETLLASIDELATKAI 61

DB 1 acnsgkdgtansadesvkgpnltelskltidnavllavkeveallssidelaakal 60

DB 62 GKRI-GNNGLEANSKNTSLSGAVASIDLAELKLVKNEELKEKIDPAKOCSTFETNK 120

DB 61 gkrlhngldteyhnngsllegayaistlikqldlgikneglkeidaakksetftnk 120

DB 121 LKSEHAVLGLDNLTDNAORAILKHKANKDGAELKLFKAVENTSKAAOTLKNAVKE 180

DB 121 lkehthdlygkgytdadakeallkngtktgaeelgkfesvevlskaakemlansvke 180

DB 181 LSPPIVHG-----NNSGKDGNTSANSADSVKGNPLTEISKTTESNAVVLAVKEIETLL 235

DB 181 lspvvaespaamvnnsgkdgtansadesvkgpnltelskltidnavllavkeveall 240

DB 236 ASTDELATRAIKKIQONGSLAVEAGHNGTLLAGAVTISKLTQKIDGLKNSKLEKIE 295

DB 241 tsidela-kalgkklkndvslneadngslisgaylislkksaalkdsgelekaele 299

DB 296 NAKKSEDFTKLEGEHAQIGIENVTDENAKKAILTTDAKDKGAELKLFKAVENTLAK 355

DB 300 kakkseeftaklgehtdlygkgytdadnakkalkltmndktgadelleklfesvknlsk 359

DB 356 AAKEMLANSVKEL 368

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

DB 360 aakemltnsvkel 372

PA (UTNY ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29032.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 to Lyme disease -

PS Claim 43; Page 123; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.

XX Sequence 398 AA;

Query Match 69.3%; Score 1245.5; DB 22; Length 398;  
 Best Local Similarity 70.8%; Pred. No. 3,6e-73;  
 Matches 264; Conservative 43; Mismatches 59; Indels 7; Gaps 3;

DB 2 ACNNSGKDGNSANSADSVKGNPLTEISKTTESNAVVLAVKEVETLLASIDELATKAI 61

DB 25 acnsgkdgtansadesvkgpnltelskltidnavllavkeveallssidelaakal 84

DB 62 GKRI-GNNGLEANSKNTSLSGAVASIDLAELKLVKNEELKEKIDPAKOCSTFETNK 120

DB 85 gkrlhngldteyhnngsllegayaistlikqldlgikneglkeidaakksetftnk 144

DB 121 LKSEHAVLGLDNLTDNAORAILKHKANKDGAELKLFKAVENTSKAAOTLKNAVKE 180

DB 145 lkehthdlygkgytdadakeallkngtktgaeelgkfesvevlskaakemlansvke 204

DB 181 LSPPIVHG-----NNSGKDGNTSANSADSVKGNPLTEISKTTESNAVVLAVKEIETLL 235

DB 205 ltpvvaespaamvnnsgkdgtansadesvkgpnltelskltidnavllavkeveall 264

DB 236 ASTDELATRAIKKIQONGSLAVEAGHNGTLLAGAVTISKLTQKIDGLKNSKLEKIE 295

DB 265 tsidela-kalgkklkndvslneadngslisgaylislkksaalkdsgelekaele 323

DB 296 NAKKSEDFTKLEGEHAQIGIENVTDENAKKAILTTDAKDKGAELKLFKAVENTLAK 355

DB 324 kakkseeftaklgehtdlygkgytdadnakkalkltmndktgadelleklfesvknlsk 383

DB 356 AAKEMLANSVKEL 368

DB 384 aakemltnsvkel 396

Search completed: March 18, 2002, 09:54:34  
 Job time: 332 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:45 ; Search time 621.2 Seconds

(without alignments)  
164,485 Million cell updates/sec

Title: US-09-596-746A-42

Sequence: 1798

Sequence: 1 MACNNSGKDGASNSADES.....AVENLAKRAKMLANSVKEL 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Pending Patents, AA, Main:\*

1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1798	100.0	368	19	US-09-596-746A-42
2	1793	99.7	367	19	US-09-596-746A-42
3	1790	99.6	391	19	US-09-596-746A-74
4	1790	99.6	392	19	US-09-596-746A-74
5	1562.5	86.9	369	19	US-09-596-746A-40
6	1557.5	86.6	368	19	US-09-596-746A-40
7	1554.5	86.5	392	19	US-09-596-746-72
8	1554.5	86.5	393	19	US-09-596-746A-72
9	1538	85.5	368	19	US-09-596-746-38

10	1538	85.5	368	19	US-09-596-746A-38	Sequence 38, Appl
11	1530	85.1	391	19	US-09-596-746-76	Sequence 76, Appl
12	1530	85.1	392	19	US-09-596-746A-76	Sequence 76, Appl
13	1489	82.8	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1485.5	82.6	377	19	US-09-596-746-30	Sequence 30, Appl
15	1485.5	82.6	377	19	US-09-596-746A-30	Sequence 30, Appl
16	1485	82.6	370	19	US-09-596-746A-36	Sequence 36, Appl
17	1484	82.5	373	19	US-09-596-746-26	Sequence 26, Appl
18	1481	82.4	397	19	US-09-596-746-62	Sequence 62, Appl
19	1481	82.4	398	19	US-09-596-746A-62	Sequence 62, Appl
20	1480	82.3	369	19	US-09-596-746-36	Sequence 36, Appl
21	1479.5	82.3	400	19	US-09-596-746-82	Sequence 82, Appl
22	1479.5	82.3	401	19	US-09-596-746A-82	Sequence 82, Appl
23	1477	82.1	393	19	US-09-596-746-66	Sequence 66, Appl
24	1477	82.1	394	19	US-09-596-746A-66	Sequence 66, Appl
25	1253.5	69.7	375	19	US-09-596-746A-24	Sequence 24, Appl
26	1250	69.5	378	19	US-09-596-746-28	Sequence 28, Appl
27	1250	69.5	378	19	US-09-596-746A-28	Sequence 28, Appl
28	1249.5	69.5	371	19	US-09-596-746A-34	Sequence 34, Appl
29	1248.5	69.4	374	19	US-09-596-746-24	Sequence 24, Appl
30	1245.5	69.3	378	19	US-09-596-746-60	Sequence 60, Appl
31	1245.5	69.3	399	19	US-09-596-746A-60	Sequence 60, Appl
32	1244.5	69.2	370	19	US-09-596-746-34	Sequence 34, Appl
33	1244	69.2	401	19	US-09-596-746A-80	Sequence 80, Appl
34	1244	69.2	402	19	US-09-596-746A-80	Sequence 80, Appl
35	1241.5	69.0	394	19	US-09-596-746-64	Sequence 64, Appl
36	1241.5	69.0	395	19	US-09-596-746A-64	Sequence 64, Appl
37	1225	68.1	370	19	US-09-596-746A-32	Sequence 32, Appl
38	1220	67.9	369	19	US-09-596-746-32	Sequence 32, Appl
39	1217	67.7	393	19	US-09-596-746-68	Sequence 68, Appl
40	1217	67.7	394	19	US-09-596-746A-68	Sequence 68, Appl
41	1167.5	64.9	360	19	US-09-596-746-52	Sequence 52, Appl
42	1167.5	64.9	360	19	US-09-596-746A-52	Sequence 52, Appl
43	1164.5	64.8	385	19	US-09-596-746A-56	Sequence 56, Appl
44	1159.5	64.5	384	19	US-09-596-746-56	Sequence 56, Appl
45	1156.5	64.3	408	19	US-09-596-746-78	Sequence 78, Appl

#### ALIGNMENTS

RESULT 1  
US-09-596-746A-42  
; Sequence 42, Application US/09596746A  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seinoist, Gerald  
; APPLICANT: Dykhulzen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746A  
; CURRENT FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746A-42

Query Match 100.0%; Score 1798; DB 19; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5.6e-124;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MACNNSGKDGASNSADESVKGNLFEISKITTESNAVLAWEVETLLASIDELATKA 60

```
Db 1 MACNNSGKDNANASANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDELATKA 60
QY 61 IGRKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 120
Db 61 IGRKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 120
QY 121 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 180
Db 121 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 180
QY 181 LTSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDE 240
Db 181 LTSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDE 240
QY 241 LATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKC 300
Db 241 LATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKC 300
QY 301 SEDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 360
Db 301 SEDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 360
QY 361 LANSVKEL 368
Db 361 LANSVKEL 368
```

```
RESULT 2
US-09-596-746-42
; Sequence 42, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ospc Chimera
US-09-596-746-42
```

Query Match 99.7%; Score 1793; DB 19; Length 367;  
Best Local Similarity 100.0%; Pred. No. 1,3e-123;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 ACNNSGKDNANASANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDELATKA 61
Db 1 ACNNSGKDNANASANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDELATKA 60
QY 62 GKRIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 121
Db 62 GKRIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 120
QY 122 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 181
Db 122 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 180
QY 182 TSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDEL 241
Db 182 TSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDEL 241
```

```
Db 181 TSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDEL 240
QY 242 ATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKCS 301
Db 242 ATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKCS 300
QY 302 EDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 361
Db 302 EDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 360
QY 362 ANSVKEL 368
Db 362 ANSVKEL 367
```

```
RESULT 3
US-09-596-746-74
; Sequence 74, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 391
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-74
```

Query Match 99.6%; Score 1790; DB 19; Length 391;  
Best Local Similarity 99.7%; Pred. No. 2.4e-123;  
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 ACNNSGKDNANASANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDELATKA 61
Db 25 SCNNSGKDNANASANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDELATKA 84
QY 62 GKRIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 121
Db 85 GKRIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNK 144
QY 122 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 181
Db 145 KSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKEKLFKAVENLSKAADOTLKNAVKEL 204
QY 182 TSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDEL 241
Db 205 TSPIVHGNNSGKDNNTSANSADSVKGPMLTEISKITTESNAVLAWEETLLASIDEL 264
QY 242 ATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKCS 301
Db 265 ATKAIGKRIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKLEKIEENAKKCS 324
QY 302 EDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 361
Db 325 EDEFTKLEGEHQAOLGIENTVDENAKKAILITDAARDKGAELKEKLFKAVENLAKAKEM 384
QY 362 ANSVKEL 368
Db 385 ANSVKEL 391
```

```
RESULT 4
US-09-596-746A-74
; Sequence 74, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoost, Gerald
; APPLICANT: Dykhuisen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746A-74

Query Match          99.6%; Score 1790; DB 19; Length 392;
Best Local Similarity 99.7%; Pred. No. 2.4e-123;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACNNSGKDNANASADESVKGNLTETSKKTESNAVYLAKEVEETLLASIDELATKAI 61
DB 26 SCNNSGKDNANASADESVKGNLTETSKKTESNAVYLAKEVEETLLASIDELATKAI 85
OY 62 GKKGNGLEANSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 121
DB 86 GKKGNGLEANSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 145
OY 122 KSEHVAVLGIDNLTDDNAQRAILKKHANKDGAELKLFKAVENLSKAODTLKNAVKEL 181
DB 146 KSEHVAVLGIDNLTDDNAQRAILKKHANKDGAELKLFKAVENLSKAODTLKNAVKEL 205
OY 182 TSPVHGNNSGKDNANSADSVKGNLTETSKKTESNAVYLAKEVEETLLASIDEL 241
DB 206 TSPVHGNNSGKDNANSADSVKGNLTETSKKTESNAVYLAKEVEETLLASIDEL 265
OY 242 ATKAIKKTIQONGGLAVEAGHNGTLLAGAVTISKLTITQKLDGKSEKLEKTEENAKKS 301
DB 266 ATKAIKKTIQONGGLAVEAGHNGTLLAGAVTISKLTITQKLDGKSEKLEKTEENAKKS 325
OY 302 EDTTKLEGEHNOIGTENTDENAKKAILITDAKDKGAELKLFKAVENLSKAODTLKNAVKEL 361
DB 326 EDTTKLEGEHNOIGTENTDENAKKAILITDAKDKGAELKLFKAVENLSKAODTLKNAVKEL 385
OY 362 ANSVKEL 368
DB 386 ANSVKEL 392

RESULT 5
US-09-596-746A-40
; Sequence 40, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoost, Gerald
; APPLICANT: Dykhuisen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-40

Query Match          86.9%; Score 1562.5; DB 19; Length 369;
Best Local Similarity 87.5%; Pred. No. 1.3e-106;
Matches 322; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNNSGKDNANASADESVKGNLTETSKKTESNAVYLAKEVEETLLASIDELATKA 60
DB 1 MACNNSGKDNANASADESVKGNLTETSKKTESNAVYLAKEVEETLLASIDELATKA 60
OY 61 IGKKGNGLEANSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 120
DB 61 IGKKGNGLEANSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 120
OY 121 LKSEHVAVLGIDNLTDDNAQRAILKKHANKDGAELKLFKAVENLSKAODTLKNAVKEL 180
DB 121 LKSEHVAVLGIDNLTDDNAQRAILKKHANKDGAELKLFKAVENLSKAODTLKNAVKEL 180
OY 181 TSPVHGNNSGKDNANSADSVKGNLTETSKKTESNAVYLAKEVEETLLASIDEL 240
DB 181 TSPVHGNNSGKDNANSADSVKGNLTETSKKTESNAVYLAKEVEETLLASIDEL 240
OY 241 LATKAIKKTIQONGGLAVEAGHNGTLLAGAVTISKLTITQKLDGKSEKLEKTEENAKKC 300
DB 241 LA-KAIKKTIQNDVSLDNEHNGSLISGAYLISMLTKRISAKSGELKAELEKAKKC 299
OY 301 SEPTTKLEGEHNOIGTENTDENAKKAILITDAKDKGAELKLFKAVENLSKAODTLKNAVKEL 360
DB 300 SEPTTKLEGEHNOIGTENTDENAKKAILITDAKDKGAELKLFKAVENLSKAODTLKNAVKEL 359
OY 361 ANSVKEL 368
DB 360 LTNSVKEL 367

RESULT 6
US-09-596-746-40
; Sequence 40, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoost, Gerald
; APPLICANT: Dykhuisen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIORITY FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OspC Chimera
US-09-596-746-40

Query Match          86.6%; Score 1557.5; DB 19; Length 368;
```

Best Local Similarity 87.5%; Pred. No. 3e-106;  
Matches 321; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 61
DB 1 ACNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 60
OY 62 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 121
DB 61 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 120
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 181
DB 121 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 180
OY 182 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 241
DB 181 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 240
OY 242 ATKAIGKTIQONGGLAVEGHNGTLLAGAYTTSKLTORLKLKNSKLEKIEENAKKS 301
DB 241 A-KAIGKTIKNDVSLDNEADHNGSLISGAYILSNLTTKISAINKSGELKAEIEKAKKS 299
OY 302 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 361
DB 300 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 359
OY 362 ANSVKEL 368
DB 360 TNSVKEL 366

RESULT 7
US-09-596-746-72
; Sequence 72, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140.042
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 392
; TYPE: PRN
; ORGANISM: ospc Chimera
US-09-596-746-72

```

Query Match 86.5%; Score 1554.5; DB 19; Length 392;  
Best Local Similarity 87.2%; Pred. No. 5.4e-106;  
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 61
DB 25 SCNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 84
OY 62 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 121
DB 85 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 144
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 181
DB 145 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 204

```

```

OY 182 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 241
DB 205 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 264
OY 242 ATKAIGKTIQONGGLAVEGHNGTLLAGAYTTSKLTORLKLKNSKLEKIEENAKKS 301
DB 265 A-KAIGKTIKNDVSLDNEADHNGSLISGAYILSNLTTKISAINKSGELKAEIEKAKKS 323
OY 302 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 361
DB 324 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 383
OY 362 ANSVKEL 368
DB 384 TNSVKEL 390

```

```

RESULT 8
US-09-596-746a-72
; Sequence 72, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140.042
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 393
; TYPE: PRN
; ORGANISM: ospc Chimera
US-09-596-746a-72

```

Query Match 86.5%; Score 1554.5; DB 19; Length 393;  
Best Local Similarity 87.2%; Pred. No. 5.5e-106;  
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```

OY 2 ACNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 61
DB 26 SCNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDELATKAI 85
OY 62 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 121
DB 86 GKRIKGNNGLEANSKNTSLISGAYAISDLIAEKLNVLKNELEKIDTAKOCSTETNTKL 145
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 181
DB 146 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKRAVENLSKAADOTLKNAVKEL 205
OY 182 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 241
DB 206 TSPYVGNNSGKDNASANSADSVKGNLTKTESNAVLAKEVEETLLASIDEL 265
OY 242 ATKAIGKTIQONGGLAVEGHNGTLLAGAYTTSKLTORLKLKNSKLEKIEENAKKS 301
DB 266 A-KAIGKTIKNDVSLDNEADHNGSLISGAYILSNLTTKISAINKSGELKAEIEKAKKS 324
OY 302 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 361
DB 325 EEDTKLGEHQAOLGIENTVDENAKKAILITDAKDGAELKRAVENLSKAADOTLKNAVKEL 384
OY 362 ANSVKEL 368
DB 385 TNSVKEL 391

```



```

RESULT 9
US-09-596-746-38
; Sequence 38, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OSpC Chimera
US-09-596-746-38

```

```

Query Match      85.5%: Score 1538: DB 19: Length 368:
Best Local Similarity 86.7%: Pred. No. 8.1e-105:
Matches 319: Conservative 20; Mismatches 27; Indels 2; Gaps 2:

```

```

QY 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
QY 61 ICKKIGNGLEANSKNTSLSGAVAYSIDLIAEKLAVLNKNEELKEKIDTAKOCSTEFNTK 120
DB 61 ICKKIGNGLEANSKNTSLSGAVAYSIDLIAEKLAVLNKNEELKEKIDTAKOCSTEFNTK 120
QY 121 LKSEHAVLGLDNLTDNAORAILKHKHANKDGAELKLEKFAVENLSKAADPTLKNVKE 180
DB 121 LKSEHAVLGLDNLTDNAORAILKHKHANKDGAELKLEKFAVENLSKAADPTLKNVKE 180
QY 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
DB 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
QY 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKITLTKDGLKNSKIEKIEENAKKC 300
DB 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKITLTKDGLKNSKIEKIEENAKKC 300
QY 301 SEDFTKKLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADPT 360
DB 301 SEDFTKKLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADPT 360
QY 361 LANSVKEL 368
DB 359 LKNAVKEL 366

```

```

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OSpC Chimera
US-09-596-746A-38

Query Match      85.5%: Score 1538: DB 19: Length 368:
Best Local Similarity 86.7%: Pred. No. 8.1e-105:
Matches 319: Conservative 20; Mismatches 27; Indels 2; Gaps 2:

QY 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
QY 61 ICKKIGNGLEANSKNTSLSGAVAYSIDLIAEKLAVLNKNEELKEKIDTAKOCSTEFNTK 120
DB 61 ICKKIGNGLEANSKNTSLSGAVAYSIDLIAEKLAVLNKNEELKEKIDTAKOCSTEFNTK 120
QY 121 LKSEHAVLGLDNLTDNAORAILKHKHANKDGAELKLEKFAVENLSKAADPTLKNVKE 180
DB 121 LKSEHAVLGLDNLTDNAORAILKHKHANKDGAELKLEKFAVENLSKAADPTLKNVKE 180
QY 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
DB 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
QY 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKITLTKDGLKNSKIEKIEENAKKC 300
DB 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKITLTKDGLKNSKIEKIEENAKKC 300
QY 301 SEDFTKKLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADPT 360
DB 301 SEDFTKKLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADPT 360
QY 361 LANSVKEL 368
DB 359 LKNAVKEL 366

RESULT 11
US-09-596-746-76
; Sequence 76, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 391
; TYPE: PRT
; ORGANISM: OSpC Chimera
US-09-596-746-76

```

Query Match 85.1%; Score 1530; DB 19; Length 391;  
Best Local Similarity 86.4%; Pred. No. 3.5e-104;  
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

QY 2 ACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 61  
DB 25 SCNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 84  
QY 62 GKRTGNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTNKL 121  
DB 85 GKRTGNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTNKL 144  
QY 122 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 181  
DB 145 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 204  
QY 182 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 241  
DB 205 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 264  
QY 242 ATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTITOKLDGKNSKLEKIEENAKKCS 301  
DB 265 ATKAIGKTI-GNNGLEANSKNTSLSGAVALSDLAETKLVNKN-ELKEKIDTAKOCS 322  
QY 302 EDPFKLGEHAOLGIENVTDENAKKAILITDPAKDGAEELEKFKAVENLSKAADTL 361  
DB 323 TEFTNKLKSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTL 382  
QY 362 ANSVKEL 368  
DB 383 KNAVKEL 389

RESULT 12  
US-09-596-746a-76  
; Sequence 76, Application US/09596746a  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seimost, Gerald  
; APPLICANT: Dykhuizen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746a  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 76  
; LENGTH: 392  
; TYPE: PRT  
; ORGANISM: ospc Chimera  
US-09-596-746a-76

Query Match 85.1%; Score 1530; DB 19; Length 392;  
Best Local Similarity 86.4%; Pred. No. 3.5e-104;  
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

QY 2 ACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 61  
DB 26 SCNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 85  
QY 62 GKRTGNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTNKL 121  
DB 86 GKRTGNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTNKL 145  
QY 122 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 181

DB 146 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 205  
QY 182 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 241  
DB 206 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 265  
QY 242 ATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTITOKLDGKNSKLEKIEENAKKCS 301  
DB 266 ATKAIGKTI-GNNGLEANSKNTSLSGAVALSDLAETKLVNKN-ELKEKIDTAKOCS 323  
QY 302 EDPFKLGEHAOLGIENVTDENAKKAILITDPAKDGAEELEKFKAVENLSKAADTL 361  
DB 324 TEFTNKLKSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTL 383  
QY 362 ANSVKEL 368  
DB 384 KNAVKEL 390

RESULT 13  
US-09-596-746a-26  
; Sequence 26, Application US/09596746a  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seimost, Gerald  
; APPLICANT: Dykhuizen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746a  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Ospc Chimera  
US-09-596-746a-26

Query Match 82.8%; Score 1489; DB 19; Length 374;  
Best Local Similarity 83.4%; Pred. No. 3.4e-101;  
Matches 312; Conservative 22; Mismatches 34; Indels 6; Gaps 2;

QY 1 MACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKA 60  
DB 1 MACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKA 60  
QY 61 IGKRTI-GNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTN 119  
DB 61 IGKRTI-GNNGLEANSKNTSLSGAVALSDLAETKLVNKNELKEKIDTAKOCSTEEFTN 120  
QY 120 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 179  
DB 121 KSEHAVLGIDNLTDDNAORAILKKHANKDKGAEELEKFKAVENLSKAADTLKNAVKEL 180  
QY 180 ELTSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 234  
DB 181 ELTSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 240  
QY 235 LMSIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTITOKLDGKNSKLEKIE 294  
DB 241 LMSIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTITOKLDGKNSKLEKIE 300  
QY 295 ENAKKSEDPFKLGEHAOLGIENVTDENAKKAILITDPAKDGAEELEKFKAVENLSKA 354  
DB 301 ENAKKSEDPFKLGEHAOLGIENVTDENAKKAILITDPAKDGAEELEKFKAVENLSKA 360

OY 355 KAKEMLANSVKEL 368  
|  
Db 361 KAKEMLANSVKEL 374

## RESULT 14

US-09-596-746-30  
; Sequence 30, Application US/09596746  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seinoost, Gerald  
; APPLICANT: Dykhuizen, Daniel  
; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 377  
; TYPE: PRF  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746-30

Query Match 82.6%; Score 1485.5; DB 19; Length 377;  
Best Local Similarity 82.5%; Pred. No. 6.2e-101;  
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 1 MACNNSGKDGNANSADSVKGPMLTEISKITSSNAVLAKEVEILLASIDELATKA 60  
|  
Db 1 MACNNSGKDGNANSADSVKGPMLTEISKITSSNAVLAKEVEILLASIDELATKA 60  
OY 61 ICKKI-GNNGLEBANOSKNTSLSGAVAISDLAEKLVNKLNELEKIDTAKOCSTEFN 119  
|  
Db 61 ICKKI-HONNGDLENNHNSLGAFAISTLIKOKLDGLKNGLEKIDAKKCSFTFN 120  
OY 120 KIKSEHAVLGIDNLTDDNARAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAV 179  
|  
Db 121 KIKSEHAVLGIDNLTDDNARAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAV 180  
OY 180 ELTSPVH-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 231  
|  
Db 181 ELTSPVH-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 240  
OY 232 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKLITOKLDLKNSEK 291  
|  
Db 241 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKLITOKLDLKNSEK 300  
OY 292 EKIEAKKCSDEFTKKLEGEHAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVE 351  
|  
Db 301 EKIEAKKCSDEFTKKLEGEHAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVE 360  
OY 352 NLAKAKEMLANSVKEL 368  
|  
Db 361 NLAKAKEMLANSVKEL 377

## RESULT 15

US-09-596-746a-30  
; Sequence 30, Application US/09596746A  
; GENERAL INFORMATION:  
; APPLICANT: Dattwyler, Raymond J.  
; APPLICANT: Seinoost, Gerald  
; APPLICANT: Dykhuizen, Daniel

; APPLICANT: Luft, Benjamin J.  
; APPLICANT: Maria J.C. Gomes-Solecki  
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
; FILE REFERENCE: 2631.1002-001  
; CURRENT APPLICATION NUMBER: US/09/596,746A  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: US 60/140,042  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 377  
; TYPE: PRF  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: OspC Chimera  
US-09-596-746a-30

Query Match 82.6%; Score 1485.5; DB 19; Length 377;  
Best Local Similarity 82.5%; Pred. No. 6.2e-101;  
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 1 MACNNSGKDGNANSADSVKGPMLTEISKITSSNAVLAKEVEILLASIDELATKA 60  
|  
Db 1 MACNNSGKDGNANSADSVKGPMLTEISKITSSNAVLAKEVEILLASIDELATKA 60  
OY 61 ICKKI-GNNGLEBANOSKNTSLSGAVAISDLAEKLVNKLNELEKIDTAKOCSTEFN 119  
|  
Db 61 ICKKI-HONNGDLENNHNSLGAFAISTLIKOKLDGLKNGLEKIDAKKCSFTFN 120  
OY 120 KIKSEHAVLGIDNLTDDNARAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAV 179  
|  
Db 121 KIKSEHAVLGIDNLTDDNARAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAV 180  
OY 180 ELTSPVH-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 231  
|  
Db 181 ELTSPVH-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 240  
OY 232 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKLITOKLDLKNSEK 291  
|  
Db 241 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKLITOKLDLKNSEK 300  
OY 292 EKIEAKKCSDEFTKKLEGEHAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVE 351  
|  
Db 301 EKIEAKKCSDEFTKKLEGEHAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVE 360  
OY 352 NLAKAKEMLANSVKEL 368  
|  
Db 361 NLAKAKEMLANSVKEL 377

Search completed: March 18, 2002, 10:08:46  
Job time: 974 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:35 ; Search time 55.5 Seconds

(without alignments)  
149.211 Million cell updates/sec

Title: US-09-596-746A-42

Perfect score: 1798  
Sequence: 1 MACNNSGKDDGNASNADES.....AVENTLAKAKEMLANSVKEL 368Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	699	38.9	588	4	US-08-235-836C-122 Sequence 122, App
2	698.5	38.8	466	4	US-08-235-836C-107 Sequence 107, App
3	665	37.0	466	4	US-08-235-836C-110 Sequence 110, App
4	657.5	36.6	212	1	US-08-158-353-2 Sequence 2, Appl
5	644.5	35.8	209	4	US-09-196-293-15 Sequence 15, Appl
6	644.5	35.8	210	1	US-08-158-353-3 Sequence 3, Appl
7	639.5	35.6	210	4	US-08-235-836C-30 Sequence 30, Appl
8	637.5	35.5	210	4	US-08-209-603E-15 Sequence 15, Appl
9	600.5	33.4	212	1	US-08-031-293-2 Sequence 2, Appl
10	600.5	33.4	212	4	US-07-903-580-2 Sequence 2, Appl
11	588.5	32.7	212	1	US-08-158-353-4 Sequence 4, Appl
12	583.5	32.5	212	4	US-08-235-836C-34 Sequence 34, Appl
13	578.5	32.2	212	4	US-09-196-293-11 Sequence 11, Appl
14	578.5	32.2	212	4	US-08-209-603E-11 Sequence 11, Appl
15	563	31.3	207	4	US-08-235-836C-36 Sequence 36, Appl
16	553	30.8	209	4	US-08-235-836C-32 Sequence 32, Appl
17	168.5	9.4	1196	2	US-08-144-121-4 Sequence 4, Appl
18	168.5	9.4	1196	2	US-08-735-893-4 Sequence 4, Appl
19	152.5	8.5	630	4	US-08-973-462-9 Sequence 9, Appl
20	151.5	8.4	1181	1	US-08-053-614-4 Sequence 4, Appl
21	151.5	8.4	1181	1	US-08-316-397B-4 Sequence 4, Appl
22	151.5	8.4	1181	2	US-09-034-306-4 Sequence 4, Appl
23	151.5	8.4	1181	2	US-09-259-437-4 Sequence 4, Appl
24	151.5	8.4	1181	5	PCR-US93-09782-4 Sequence 4, Appl
25	151.5	8.4	2482	1	US-08-328-254-6 Sequence 6, Appl
26	150.5	8.4	3248	1	US-08-353-700-1 Sequence 1, Appl
27	150.5	8.4	3248	5	PCR-US95-16216-1 Sequence 1, Appl

28	146.5	8.1	1164	4	US-08-923-992A-2 Sequence 2, Appl
29	146.5	8.1	1786	4	US-08-973-462-8 Sequence 8, Appl
30	145	8.1	1147	3	US-08-470-260-5 Sequence 5, Appl
31	145	8.1	1147	3	US-08-471-491-5 Sequence 5, Appl
32	145	8.1	1147	4	US-08-466-662-5 Sequence 5, Appl
33	145	8.1	3289	2	US-08-477-451-2 Sequence 2, Appl
34	142	7.9	688	3	US-09-141-047-8 Sequence 8, Appl
35	141.5	7.9	3111	2	US-08-460-309-4 Sequence 4, Appl
36	141.5	7.9	3111	2	US-08-125-077-4 Sequence 4, Appl
37	141	7.8	2101	1	US-08-466-390-4 Sequence 4, Appl
38	141	7.8	2101	1	US-08-470-950-4 Sequence 4, Appl
39	141	7.8	2101	1	US-08-467-781-4 Sequence 4, Appl
40	141	7.8	2101	1	US-08-195-487-4 Sequence 4, Appl
41	141	7.8	2101	2	US-08-483-924-4 Sequence 4, Appl
42	141	7.8	2101	4	US-09-452-294-1 Sequence 4, Appl
43	141	7.8	2101	5	PCR-US93-06160-4 Sequence 4, Appl
44	138.5	7.7	1128	4	US-08-923-992A-6 Sequence 6, Appl
45	136.5	7.6	679	4	US-08-913-942-15 Sequence 15, Appl

## ALIGNMENTS

RESULT 1  
US-08-235-836C-122  
Sequence 122, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 122:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 588 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-122

Query Match 38.9%; Score 699; DB 4; Length 588;  
Best Local Similarity 41.5%; Pred. No. 5.1e-45;  
Matches 195; Conservative 49; Mismatches 104; Indels 122; Gaps 17;  
QY 7 GKGMSANSNDSVKG-----PMLTETSKITTESNAVLAIVKEVER-----LNASID 54

Db 9 GSQKENDLNEDSSKSHQNAKQDLPAYTEDESVSLFNGNKIIVSKKKNSSGKYDRLATID 68  
OY 55 ELATAIKGKIGNNG--LEANQ-SKNTSLSGAIAISDLIAEKLVKNELEKKEIDTA 110  
Db 69 OVELGTGSDK--NMSGTLEGSKPKSKVLTIVSADLVTLVLEAPDA-SNOKISSKV-TK 124  
OY 111 KQCS-TEFT---NKL-----KSEHAVLGLODLT/D-DNAQRAI----- 142  
Db 125 KQGSTTEETLRANKLDSKRLIRSNCTTLEYSTITADATKAVETLKSIRKEGSLVYCK 184  
OY 143 -----LKKHAKND----- 150  
Db 185 TTVEIKEGTVLKRLEIKEDGKVKVFLNDTAGSNKKTGKWEDESTSTLTISADSKTKRDLVF 244  
OY 151 -----KGAELKLEKFAVENISKAQDPLKNAVKELTPIYHGNSSGDKGNTS 198  
Db 245 LTDGTTVOQYNTAGTSLEGSASEIKNLSE-----LKNALK--GHPM--GNSSGDKGNTS 295  
OY 199 ANSADESVKGNPLTEISKRTTESNAVLAKEIETLLASIDELATKAIKGIQNGGLAV 258  
Db 296 ANSADESVKGNPLTEISKRTTESNAVLAKEIETLLASIDELATKAIKGIQNGGLAV 355  
OY 239 EAGHNGTLAAGYTSKLTOKLDGLKNSKIEKIEKAKCSEPTFKLGEHAQIGIE 318  
Db 356 EYNHNGSLAGRYAISTLIKOKLDGLKN-EGLEKIDAKKCEFTFKLKEKHTDLGKE 414  
OY 319 NTDENAKKALITDAKOKGAELKLEKFAVENISKAQDPLKNAVKEL 368  
Db 415 GVTADAKKALITKNGTKTKGAELKLEKFAVENISKAQDPLKNAVKEL 464

RESULT 2  
US-08-235-836C-107  
; Sequence 107, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Luft, Benjamin J.  
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235.836C  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 466 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-08-235-836C-107

Query Match 38.8%; Score 698.5; DB 4; Length 466;  
Best Local Similarity 47.7%; Pred. No. 4,1e-45;  
Matches 177; Conservative 43; Mismatches 114; Indels 37; Gaps 8;

OY 19 ESVKGNPLTEISKRTTESNAVLAKEIETLLASIDELATKAIKGIQNGGLAV 77  
Db 100 EYFKEDGKTLVSKKWTSDKDSSTEEKFNEKEGYS-EKITRADGRLETYTGKSGSGSKA 158  
OY 78 TSLSGAAYAISDLIAEKLVKNE--ELKEKIDAKCSEFTFKLGEHAQIGIE 134  
Db 159 KEVLKGYVLEGTLPFAKETTLVKEGTVLSKNIKSGEVSVELNDPTDSS--AATKRTAAW 216  
OY 135 DDNAORALKKHAKNDK-----GALEKLEKFAVENISKAQDPLKNA 177  
Db 217 NDSTSTLTISADSKTKVLPVLDGTTVOQYNTAGTSLEGSASEIKNLSE-----LKN 271  
OY 178 VKELTSPYVHGNSSGDKGNTSANSADSVKGNPLTEISKRTTESNAVLAKEIETLLAS 237  
Db 272 LK-----MANNSSGDKGNTSANSADSVKGNPLTEISKRTTESNAVLAKEIETLLAS 324  
OY 238 IDELATKAIKGIQNGGLAVAGHNGTLAAGYTSKLTOKLDGLKNSKIEKIEKAKCSEPTFKL 297  
Db 325 IDELATKAIKGIQNGGLAVAGHNGTLAAGYTSKLTOKLDGLKN-EGLEKIDAKKCEFTFKL 383  
OY 298 KKCSEFTFKLGEHAQIGIEVNDENAKKALITDAKOKGAELKLEKFAVENISKAQDPLKNA 357  
Db 384 KKCSEFTFKLGEHAQIGIEVNDENAKKALITDAKOKGAELKLEKFAVENISKAQDPLKNA 443  
OY 358 KEMLANSVKEL 368  
Db 444 KEMLANSVKEL 454

RESULT 3  
US-08-235-836C-110  
; Sequence 110, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Luft, Benjamin J.  
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235.836C  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 37.0%; Score 665; DB 4; Length 466;  
Best Local Similarity 46.9%; Pred. No. 1,3e-42;  
Matches 165; Conservative 47; Mismatches 104; Indels 36; Gaps 8;

QY 1 MACNNSGKGNASANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 60  
DB 17 ISCNNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 76  
QY 61 IGRKI-GNNGLEANOSKNTSLSGAAYAISDLIAEKLNVLEKELIDTAKOCSTFTN 119  
DB 77 IGRKIQNGNLDPEYHNHNSLLAGRYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTN 136  
QY 120 KLRSEAVIGLNDLTDNAORALIKKHNKDKGAELKFRVAVENLSKAADOTLKNA 179  
DB 137 KLRKHTDGLKEGVTDADKKAELKTNGTKCAEELKFESEVYLSKAKEMLNKSYK 196  
QY 180 ELTSPVHGNGNSGKDGNTSA---NSADSVKGNPLTEISKTTESNAVAVLAKEVETLL 235  
DB 197 ELTSPVAVASPKPKQNVSSLDKNSVVDLPKMKVLYSKENKDKGY-----DLI 248  
QY 236 ASIDELATRAIGKIOOGL-AVEAGHNGTLA-----GAYTISKLT 278  
DB 249 ATVDKRLKGTSDKNGSGVLEGVKADKSKVKLTISDDLQGTTLVEYFKEDGKTLVSKVT 308  
QY 279 QKLDGKNSKELKEIKEMAKKCEDTKKLEGEHAQ-LGIENVTDENAKKAI 329  
DB 309 SK-DKSTEEKENKE---VSEKILITRADGTRLEYTGKISGSGKAKKVL 355

## RESULT 4

US-08-158-353-2  
Sequence 2, Application US/08158353  
Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-2

Query Match 36.6%; Score 657.5; DB 1; Length 212;  
Best Local Similarity 72.9%; Pred. No. 1,7e-42;  
Matches 140; Conservative 20; Mismatches 29; Indels 3; Gaps 2;

QY 1 MACNNSGKGNASANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 60  
DB 17 ISCNNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 76  
QY 61 IGRKI-GNNGLEANOSKNTSLSGAAYAISDLIAEKLNVLEKELIDTAKOCSTFTN 119  
DB 77 IGRKIQNGNLDPEYHNHNSLLAGRYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTN 136  
QY 120 KLRSEAVIGLNDLTDNAORALIKKHNKDKGAELKFRVAVENLSKAADOTLKNA 177  
DB 137 KLRKHTDGLKEGVTDADKKAELKTNGTKCAEELKFESEVYLSKAKEMLNKSYK 196  
QY 178 VKELTSPVHGN 189  
DB 197 VKELTSPVAVEN 208

## RESULT 5

US-09-196-293-15  
Sequence 15, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Molz, Manfred  
APPLICANT: Soutscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
TITLE OF INVENTION: burgdorferi  
FILE REFERENCE: 738,001US2  
CURRENT APPLICATION NUMBER: US/09/196,293  
EARLIER FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 209  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-15

Query Match 35.8%; Score 644.5; DB 4; Length 209;  
Best Local Similarity 75.6%; Pred. No. 1,6e-41;  
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 189 NSNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 248  
DB 20 NSNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKEVETLLASIDELATKA 79  
QY 249 KIOQNGLAVERAGHNGTLAAYTISKLTOKLDGKNSKELKEIKEMAKKCEDFTKL 308  
DB 80 KIHQNGNLDPEYHNHNSLLAGRYAISTLIKOKLDGLKN-EGLKREKIDAKKCEFTNKL 138







STATE: Virginia  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/07/903,580  
FILING DATE: 19920625  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BEHT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/131 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-903-580-2

Query Match 33.4%; Score 600.5; DB 4; Length 212;  
Best Local Similarity 70.2%; Pred. No. 3,4e-38; Indels 1; Gaps 1;  
Matches 127; Conservative 20; Mismatches 33;

QY 189 NNSGKDGNT-SANSADSVKGNLTETSKRTTESNAVLAWEIETLLASIDELATK 247  
DB 20 NNSGKGDGASSTNPADSAKGNLTETSKRTTDSNAFVLAWEIETLLASIDELATK 79  
QY 248 KRIQONGGLAVAGHNGITLACATYITSLITOKLDGLNSEKLEKIEKAKCSEDFTKK 307  
DB 80 KRIQONGGLAVAGHNGITLACATYITSLITOKLDGLNSEKLEKIEKAKCSEDFTKK 139  
QY 308 LEGEHAQGIENVTDENAKKAILITDAKDKGALEKLFRAVENLAKAEMKANSYKE 367  
DB 140 LAAGHAQGIENVTDENAKKAILITDAKDKGALEKLFRAVENLAKAEMKANSYKE 199  
QY 368 L 368  
DB 200 L 200

RESULT 11  
US-08-158-353-4  
Sequence 4, Application US/08158353  
Patent No. 5620862  
GENERAL INFORMATION:  
APPLICANT: Padula, Steven J.  
TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9340  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-4

Query Match 32.7%; Score 588.5; DB 1; Length 212;  
Best Local Similarity 60.8%; Pred. No. 2,7e-37; Indels 3; Gaps 3;  
Matches 130; Conservative 20; Mismatches 36;

QY 1 MACNNSGKDG-NASANSDESVKGNLTETSKRTTESNAVLAWEIETLLASIDELATK 59  
DB 17 ISCNNSGKGDGASSTNPADSAKGNLTETSKRTTDSNAFVLAWEIETLLASIDELATK 76  
QY 60 AIGKKI-GNNGLEANOSKNITLSGAYASIDLAELKLVLN-BELKEKIDTAKCSTEE 117  
DB 77 AIGKIDNNNNLALNONGSLAGAYASTLITEKLSKLNBELTELAKAKKCEEF 136  
QY 118 TNKTKSEHVAIGLNDLDDNQRALIKKHNKDKGALEKLFRAVENLSKAODITKNA 177  
DB 137 TNKTKSGHADGKODATDDHAKAAILKTHATTGAREFQDLFESVGLIKAAQVALTNS 196  
QY 178 VKELTSPV 186  
DB 197 VKELTSPV 205

RESULT 12  
US-08-235-836C-34  
Sequence 34, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-34

Query Match 32.5% Score 583.5; DB 4; Length 212;  
Best Local Similarity 68.3%; Pred. No. 6.5e-37;  
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNSGKDG-NASANSADSVKGNLTETSKITESNAVLAVKEVETLLASIDELAKR 59  
DB 17 ISCSNGKGDASSTNPADBSAKGPLTEISKITDSNMFVLAVKEVETLVLSIDELAKK 76  
QY 60 AIGKTI-GNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKOCSTER 117  
DB 77 AIGOKIDNNNGIALNNGNSLAGAYASTLTETKLSKLNLEELKTEIARAKKCSSEF 136  
QY 118 TTKLSEHAVLGLDNLTDNNAORAILKHKANKDGAEELEKFAVENLSKAADPTLKNA 177  
DB 137 TTKLKSGHADLGRKQDATHAKAAILKTHATTDKAKKERDLFESVEGLKKAQVALTNS 196  
QY 178 VKELTSPV 186  
DB 197 VKELTSPV 205

RESULT 13  
US-09-196-293-11  
Sequence 11, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Soutscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: burgdorferi  
CURRENT APPLICATION NUMBER: US/09/196,293  
EARLIER FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 32.2% Score 578.5; DB 4; Length 212;  
Best Local Similarity 68.3%; Pred. No. 1.5e-36;  
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

QY 1 MACNSGKDG-NASANSADSVKGNLTETSKITESNAVLAVKEVETLLASIDELAKR 59  
DB 17 ISCSNGKGVGILSTNPADBSAKGPLTEISKITDSNMFVLAVKEVETLVLSIDELAKK 76  
QY 60 AIGKTI-GNNGLEANOSKNTSLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKOCSTER 117  
DB 77 AIGOKIDNNNGIALNNGNSLAGAYASTLTETKLSKLNLEELKTEIARAKKCSSEF 136  
QY 118 TTKLSEHAVLGLDNLTDNNAORAILKHKANKDGAEELEKFAVENLSKAADPTLKNA 177  
DB 137 TTKLKSGHADLGRKQDATHAKAAILKTHATTDKAKKERDLFESVEGLKKAQVALTNS 196  
QY 178 VKELTSPV 186  
DB 197 VKELTSPV 205

RESULT 14  
US-08-209-603E-11  
Sequence 11, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILSKE, BETTINA  
APPLICANT: PREAC-MURSIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
FROM BORRELIA BURGDOFFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
DESCRIPTION: N/A  
HYPOTHETICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:

LENGTH: 20/ amino acids

Search completed: March 18, 2002, 09:55:35  
Job time: 348 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:48 ; Search time 68.77 Seconds

(without alignments)  
407.623 Million cell updates/sec

Title: US-09-596-746a-42

Sequence: 1798  
1 MACNNSCKDGNASANSADSES.....AVENLAKAKEMILANSVKEL 368Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summariesDatabase : PIR68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	901	50.1	209 2	I40142 outer surface prot
2	876	48.7	193 2	S70287 outer surface prot
3	863	48.0	191 2	I40153 outer surface prot
4	826	45.9	178 2	I40125 outer surface prot
5	823	45.8	178 2	I40104 outer surface prot
6	739	41.1	178 2	I40124 outer surface prot
7	648.5	35.1	210 2	S69927 outer surface prot
8	646	35.9	211 2	I40145 outer surface prot
9	644.5	35.8	210 2	G70218 outer surface prot
10	642	35.7	193 2	S70279 outer surface prot
11	640	35.6	193 2	S70280 outer surface prot
12	639.5	35.6	210 2	I40144 outer surface prot
13	637	35.4	211 2	S69918 outer surface prot
14	635	35.3	203 2	I40108 outer surface prot
15	627	34.9	211 2	I40268 outer surface prot
16	625	34.8	193 2	S70286 outer surface prot
17	622	34.6	209 2	S69917 outer surface prot
18	616	34.3	191 2	S70284 outer surface prot
19	613.5	34.1	194 2	S70277 outer surface prot
20	612.5	34.1	212 2	S69922 outer surface prot
21	611.5	34.0	194 2	S70268 outer surface prot
22	602	33.5	179 2	S54222 outer surface prot
23	601.5	33.5	212 2	I40143 outer surface prot
24	600.5	33.4	177 2	I40129 outer surface prot
25	600.5	33.4	212 2	S70254 outer surface prot
26	594	33.0	209 2	I40281 outer surface prot
27	591	32.9	178 2	I40106 outer surface prot
28	589	32.8	193 2	S70274 outer surface prot
29	588.5	32.7	190 2	S70261 outer surface prot

30	588.5	32.7	212 2	S20543 outer surface prot
31	586	32.6	193 2	S70276 outer surface prot
32	586	32.6	211 2	S69932 outer surface prot
33	582	32.4	211 2	I40277 outer surface prot
34	579.5	32.2	212 2	S69921 outer surface prot
35	578.5	32.2	196 2	S70290 outer surface prot
36	578.5	32.2	212 2	I40279 outer surface prot
37	577	32.1	211 2	I40278 outer surface prot
38	575	32.0	193 2	S70265 outer surface prot
39	569	31.6	207 2	I40271 outer surface prot
40	567.5	31.6	192 2	S70285 outer surface prot
41	567	31.5	200 2	I40122 outer surface prot
42	567	31.5	207 2	S69919 outer surface prot
43	567	31.5	207 2	S37727 outer surface prot
44	565.5	31.5	194 2	S70289 outer surface prot
45	564	31.4	207 2	I40276 outer surface prot

## ALIGNMENTS

RESULT 1  
I40142  
outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40142; S70283  
R:Stevenson, B.; Bockenstedt, L.R.; Barthold, S.W.  
Infect. Immun. 62, 3568-3571, 1994  
A:Title: Expression and gene sequence of outer surface protein C of Borrelia burgdorferi  
A:Reference number: I40142; MID:94314484  
A:Accession: I40142  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-209 <RES>  
A:Cross-references: EMBL:004240; NID:g2314877; PIDN:AAC45538.1; PID:g434658  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dornier, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease spirochete  
A:Reference number: S70255; MID:96296448  
A:Accession: S70283  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-209 <LIV>  
A:Cross-references: EMBL:L42894; NID:g858722; PIDN:AAB37002.1; PID:g1695219  
A:Experimental source: strain 28691  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.1%; Score 901; DB 2; Length 209;  
Best local similarity 98.9%; Pred. No. 4.4e-36;  
Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1	MACNNSCKDGNASANSADSESVKPNLTETSKTTESNAVVLAKEVETLLASIDELATKA 60
DB	17	ISCNNSCKDGNASANSADSESVKPNLTETSKTTESNAVVLAKEVETLLASIDELATKA 76
QY	61	IGKTIKNNNGLEANOSKNTSLSGAVAIISDIIAKLVNKKKEELKEKIDPAKQSTFEPTNK 120
DB	77	IGKTIKNNNGLEANOSKNTSLSGAVAIISDIIAKLVNKKKEELKEKIDPAKQSTFEPTNK 136
QY	121	LKSEHAVLIGDNLTDNAQRAILKKHANKDKGAEELEKFAVENLSKAADTLKNAVKE 180
DB	137	LKSEHAVLIGDNLTDNAQRAILKKHANKDKGAEELEKFAVENLSKAADTLKNAVKE 196
QY	181	LSPPIV 186
DB	197	LSPPIV 202
RESULT	2	

S70287

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C:Accession: S70287

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70287

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-193 &lt;LIV&gt;

A:Cross-references: EMBL:L42895; NID:9858723; PIDN:AB37003.1; PID:91695220

A:Experimental source: Strain 28354

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 48.7%; Score 876; DB 2; Length 193;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 NNSGKNTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGK 248

Db 2 NNSGKNTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGK 61

QY 249 KIQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKL 308

Db 62 KIQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKL 121

QY 309 EGEHAQGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 368

Db 122 EGEHAQGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 181

RESULT 3

I40153

outer surface protein C - Lyme disease spirochete (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 26-May-2000

C:Accession: I40153

R:Funf, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.

Infect. Immun. 62, 3213-3221, 1994

A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in I

A:Reference number: I40153; MUID:94314437

A:Accession: I40153

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 &lt;RES&gt;

A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 48.0%; Score 863; DB 2; Length 191;

Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 190 NSGKNTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGK 249

Db 1 NSGKNTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGK 60

QY 250 IQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLE 309

Db 61 IQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLE 120

QY 310 GEHAQGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 368

Db 121 GEHAQGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 179

RESULT 4

I40125

outer surface protein C - Lyme disease spirochete (strain MUI) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000

C:Accession: I40125; S54195

R:Thielsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J. Bacteriol. 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: I40104; MUID:95286481

A:Accession: I40125

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 &lt;RES&gt;

A:Cross-references: EMBL:X84779; NID:9793817; PIDN:CAA59250.1; PID:9793818

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 45.9%; Score 826; DB 2; Length 178;

Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 196 NTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGKIQONGG 255

Db 1 NTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGKIQONGG 60

QY 256 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLEGEHAOL 315

Db 61 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLEGEHAOL 120

QY 316 GIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 368

Db 121 GIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 173

RESULT 5

I40104

outer surface protein C - Lyme disease spirochete (strain 272) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000

C:Accession: I40104; S54184

R:Thielsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J. Bacteriol. 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: I40104; MUID:95286481

A:Accession: I40104

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 &lt;RES&gt;

A:Cross-references: EMBL:X84785; NID:9793787; PIDN:CAA59256.1; PID:9793788

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 45.8%; Score 823; DB 2; Length 178;

Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 196 NTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGKIQONGG 255

Db 1 NTSANSADSVKGPMLTEISKRTTESNAVLAKEIEFTLLASIDELATKAIGKIQONGG 60

QY 256 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLEGEHAOL 315

Db 61 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKEIKENAKKCEDFTKKLEGEHAOL 120

QY 316 GIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 368

Db 121 GIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 173



```

RESULT      6
140124
outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40124; S54194
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelson, B.; Lebech, A.M.; Hansen, J.; Bacteriol. 177, 3036-3044, 1995
A>Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; PMID:95286481
A:Accession: I40124
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84782; NID:g793815; PIDN:CA59253.1; PID:g793816
C:Superfamily: Lyme disease spirochete surface protein C

Query Match          41.1%; Score 648.5; DB 2; Length 178;
Best Local Similarity 89.0%; Pred. No. 1.5e+28;
Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps

196 NTSSANSADESVKGNPNLEISKITTESNAVAVLAKEIETLLASIDELATKAIGKIQQNG 255
|||||
1 NTSANSADESCKGPNLEISKITTESMAGGLAVKEIETLLASIDELATKAIGKIQQNG 60
|||||

OY    256 LAVAGHNGTLLAGAYITSKLTIQKLDGLKNSEKLEKEIKENAKKCEDEFTKLEGHAOL 315
Db    61 LAVAGHNGTSLGAYITSKLTIQKIMWINSERKITGENAKKCEDEFNKLEGHAOL 120
|||||

OY    316 GIEVNTDENMKALLITDAKKDKGAELFKLFRAVENLKAAKEMLANVKEL 368
Db    121 GTENVTDENADKALITTDAAKKDGAELAKLFRAVENLPQAAMKEMLANVKEL 173
|||||

RESULT      7
S69927
outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Variety: strain PKa
C>Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C:Accession: S69927; S72669
R:Jouris-Hajduk, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E., J. Clin. Microbiol. 33, 1860-1866, 1995
A>Title: Molecular analysis of genes encoding outer surface protein C (OspC)
A:Reference number: I40047; PMID:95395018
A:Accession: S69927
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-210 <JAU>
A:Cross-references: EMBL:X69589
R:Journal, S.
submitted to the EMBL Data Library, February 1994
A:Accession: S72669
A:Molecule type: DNA
A:Residues: 1-124,'D',126-139,'E',141-210 <JAW>
A:Cross-references: EMBL:X69589
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match          36.1%; Score 648.5; DB 2; Length 210;
Best Local Similarity 76.1%; Pred. No. 3.2e+24;
Matches 137; Conservative 14; Mismatches 28; Indels 1; Gaps

OY    189 NNSRGDNNTSSANSADESCKGNPNIETSKITTESNAVAVLAKEIETLLASIDELATKAIG 248
Db    20 NNSRGDNNTSSANSADESCKGNPNIETSKITTESNAVAVLAKEIETLLASIDELATKAIG 79
|||||

OY    249 KIQONGCLAVEAGHNGTLLAGAYITSKLTIQKLDGLKNSEKLEKEIKENAKKCEDEFTKKL 308

```

[illegible]

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49497; NID:q707092; PIDN:BA08457.1; PID:q769684  
R:Jauris-Heijde, S.; Fuchs, R.; Motz, M.; Preac-Murisc, V.; Schwab, E.; Soutschek, E.; W  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)  
A:Reference number: S37726; MUID:93268136  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212  
A:Experimental source: strain ip2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.8%; Score 644.5; DB 2; Length 210;  
Best Local Similarity 75.6%; Pred. No. 4.9e-24;  
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 189 NNSGKDNTSANSADSVKGPMLTEISKRTTESNAVLAVKEITFLASIDELATKAIG 248  
DB 20 NNSGKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVALLSIDELATAKIG 79

QY 249 KIOQNGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKKEIKETIENAKKCESEFTK 308  
DB 80 KIHQNNGLDTEHNNHNSLLAGAYTISKLTITOKLDGLN-EGIKETIDAKKCESEFTK 138

QY 309 EGEHAQIGIENVTDENAKKAILITDAA-KDGAELKLEKFAVENLAKAKEMLANSVKE 368  
DB 139 KEKHTDLKEGVTADAKKAILITKTGKGAELKLEFESVEVLAKAKEMLANSVKE 198

RESULT 10  
S70279  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223  
A:Experimental source: strain Z5015  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.7%; Score 642; DB 2; Length 193;  
Best Local Similarity 72.6%; Pred. No. 5.8e-24;  
Matches 135; Conservative 16; Mismatches 33; Indels 2; Gaps 2;

QY 3 CNNSGKDG-N-ASANSADSVKGPMLTEISKRTTESNAVLAVKEITFLASIDELATKA 61  
DB 1 CNNSGKDGNA-STNPADSVKGPMLTEISKRTTDSNVLAVKEVALLSIDELATKA 60

QY 62 GKKT-CNNCLEANOSKNTSLSGAVALSDLAETKLVNKEELKEKIDTAKCSTEFYTNK 120  
DB 61 KIHQNNGLDTEHNNHNSLLAGAYTISKLTITOKLDGKNSKKEIKETIENAKKCESEFTK 120

QY 121 LKSEHAVALGDLVLTDDNORATLKKHANKDGAELKLEKFAVENLAKAKEMLANSVKE 180  
DB 121 LKSSHTLELKGADQADDKAKAILTKHTNKGAELEKLEFVENLSKAAREMLSNSVKE 180

QY 181 LTPSPV 186  
DB 181 LTPSPV 186

RESULT 11  
S70280  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
C:Accession: S70280  
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70280  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LIV>  
A:Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226  
A:Experimental source: strain Z57  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.6%; Score 640; DB 2; Length 193;  
Best Local Similarity 76.2%; Pred. No. 7.2e-24;  
Matches 138; Conservative 16; Mismatches 25; Indels 2; Gaps 2;

QY 189 NNSGKDNTSANSADSVKGPMLTEISKRTTESNAVLAVKEITFLASIDELATKAIG 248  
DB 2 NNSGKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVALLSIDELATAKIG 60

QY 249 KIOQNGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKKEIKETIENAKKCESEFTK 308  
DB 61 KIHQNNGLDTEHNNHNSLLAGAYTISKLTITOKLDGLN-EGIKETIDAKKCESEFTK 120

QY 309 EGEHAQIGIENVTDENAKKAILITDAA-KDGAELKLEKFAVENLAKAKEMLANSVKE 367  
DB 121 KDNHAQIGIENVTDENAKKAILITDAA-KDGAELKLEKFAVENLAKAKEMLANSVKE 180

QY 368 L 368  
DB 181 L 181

RESULT 12  
I40144  
outer surface protein C precursor - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40144; S70282  
R:Stevenson, B.; Barthold, S.W.  
FEMS Microbiol. Lett. 124, 367-372, 1994  
A:Title: Expression and sequence of outer surface protein C among North American isol  
A:Reference number: I40143; MUID:95154673  
A:Accession: I40144  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664  
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995

Query Match	35.48	Score 637	DB 2	Length 211
Best Local Similarity	75.7%	Pred. No. 1,1e-23		
Matches 137	Conservative 17	Mismatches 25	Indels	Gaps 2
QY 189	NNSGKGGNTSANSADSEYKGNLTREISKITTESNAVLAVKEIFTLASIDELATRAIGK	248		
DB 20	NNSGKGGNTSANSADSEYKGNLTREISKITDSNAVLAVKEALLSIDELA-KAIGK	78		
QY 249	KIDQNGCLAVEAGHNGTLLAGATYITISKLTQKLDGLKNSKKETKEENAKKCEDPTKKL	308		
DB 79	KIKNDSIGDEAHHNESLLAGATYITISTLTITOKLSKINGSEGLEKETAIAKKCSEEFSTKL	138		
QY 309	EGEHAQGLGINVTDENAKKALITDAA-KDKGAELERIKFAKVENLAKAKKEBLANSYKE	367		

Query Match	34.9%	Score 627	DB 2	Length 211
Best Local Similarity	72.0%	Pred. NO. 3.3e-23		
Matches 131; Conservative	16;	Mismatches 33;	Indels 0;	Gaps 0

Db 20 NNSGKDGNTSANSADSVKGPMLTEISKTTDSNAVLAVKEVVLSSIDELAKKAIGK 79  
QY 249 KIOONGSLAVEAGHNGTLLAGAYTISKLIYOKLDGLNSEKLEKIEIENAKKCSSEFTKTL 308  
Db 80 KIDQNNALGTLDNHNSGLAGAYASIALITEKLSIKDSGELKAEIEKAKKCSSEFTKTL 139  
QY 309 EGEHAOLGIENVTDENKKAILITDAKDKGALELEKLFKAVENTLAKAKEMLANSVKEL 368  
Db 140 SDNOAELGIENATDDNKAAILKTHNAKDKGAELVKLSSEVAGLLKAAOAILANSVKEL 199

Search completed: March 18, 2002, 09:56:49  
Job time: 372 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:28 ; Search time 81.02 Seconds

(without alignments)  
290.158 Million cell updates/sec

Title: US-09-596-746A-42

Perfect score: 1798  
Sequence: 1 MACNNSGKDCDNASANSADES.....AVENLAKAKEMLANSVKEL 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.New:\*  
1: /cgn2\_6/ptodata/2/paa/BCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US05\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	35.6	211	6	US-09-974-992-5
2	588.5	32.7	212	6	US-09-974-992-7
3	567	31.5	207	6	US-09-974-992-3
4	176	9.8	837	6	US-09-815-242-5883
5	176	9.8	837	6	US-10-072-851-5883
6	176	9.8	837	6	US-09-815-242-13080
7	176	9.8	837	6	US-10-072-851-13080
8	170	9.5	1304	6	US-09-708-427-15045
9	170	9.5	1313	6	US-09-708-427-15044
10	169.5	9.4	2434	6	US-09-815-242-5835
11	169.5	9.4	2434	6	US-10-072-851-5835
12	169.5	9.4	6281	6	US-09-815-242-12996
13	169.5	9.4	6281	6	US-10-072-851-12996
14	168.5	9.4	1765	5	US-10-037-182-8
15	168.5	9.4	1765	5	US-09-938-275-6
16	168.5	9.4	1786	6	US-09-873-676-113
17	168.5	9.4	1786	6	US-09-561-709B-9
18	168.5	9.4	1786	6	US-10-037-182-6
19	167.5	9.3	1013	6	US-09-614-150-9159
20	163.5	9.1	1881	7	US-10-032-585-7646
21	163.5	9.1	1881	7	US-10-072-851-15590
22	162	9.0	1014	6	US-09-708-427-19883
23	162	9.0	1018	6	US-09-708-427-19882
24	162	9.0	1269	6	US-09-708-427-19881
25	161.5	9.0	730	6	US-09-708-427-3731

26	161.5	9.0	805	6	US-09-708-427-3730	Sequence 3730, Ap
27	161.5	9.0	841	6	US-09-708-427-3729	Sequence 3729, Ap
28	158.5	8.8	5795	7	US-09-815-242-12610	Sequence 12610, A
29	158.5	8.8	5795	7	US-10-072-851-12610	Sequence 12610, A
30	155.5	8.6	2056	6	US-09-614-150-4824	Sequence 4824, Ap
31	155	8.6	1038	6	US-09-708-427-32624	Sequence 32624, A
32	155	8.6	1042	6	US-09-708-427-32623	Sequence 32623, A
33	155	8.6	1043	6	US-09-708-427-32622	Sequence 32622, A
34	155	8.6	2076	6	US-09-815-242-5815	Sequence 5815, Ap
35	155	8.6	2076	7	US-10-072-851-5815	Sequence 5815, Ap
36	155	8.6	2186	6	US-09-815-242-12913	Sequence 12913, A
37	155	8.6	2186	7	US-10-072-851-12913	Sequence 12913, A
38	155	8.6	2437	6	US-09-815-242-5834	Sequence 5834, Ap
39	155	8.6	2437	7	US-10-072-851-5834	Sequence 5834, Ap
40	154.5	8.6	1144	6	US-09-708-427-15046	Sequence 15046, A
41	153	8.5	1725	5	US-10-037-182-12	Sequence 12, Appl
42	153	8.5	1786	5	US-09-938-275-7	Sequence 7, Appl
43	153	8.5	1786	5	US-10-037-182-10	Sequence 10, Appl
44	152.5	8.5	630	6	US-09-742-096-5	Sequence 5, Appl
45	152	8.5	1128	7	US-10-032-585-7627	Sequence 7627, Ap

## ALIGNMENTS

RESULT 1  
US-09-974-992-5  
; Sequence 5, Application US/0974992  
; GENERAL INFORMATION:  
; APPLICANT: Mathiesen, Marianne J.  
; APPLICANT: Thelssen, Michael  
; APPLICANT: Holm, Arne  
; APPLICANT: Ostergaard, Soren  
; TITLE OF INVENTION: Novel Ospc-derived peptide fragments  
; FILE REFERENCE: 459-666p  
; CURRENT APPLICATION NUMBER: US/09/974, 992  
; CURRENT FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: 09/180, 089  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: PCT/DK97/00203  
; PRIOR FILING DATE: 1997-05-02  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-974-992-5

Query Match 35.6%; Score 640; DB 6; Length 211;  
Best Local Similarity 76.2%; Pred. No. 8.2e-37;  
Matches 18; Conservative 16; Mismatches 25; Indels 2; Gaps 2;

QY 189 NNSGKGNSTANSANDESVKPNLFEISKRTTESNAVLAKEETILLASIDELATKRGK 248  
DB 20 NNSGKGNSTANSANDESVKPNLFEISKRTTESNAVLAKEEBALSSIDEIA-KAIGK 78  
QY 249 KIOONGLAVEACHNGTLLAGAYTISKRLYQKLDGKNSKLEKIEIENARKCEDFTKRL 308  
DB 79 KINSGSLGDEANHNESLLAGAYTISLTRYQKLSKLSGSLSEKIEIAAKKCEEFSTKL 138  
QY 309 EGEHAQGLINVTDENAKKAILTTDA-KKGAALEKLEKRAVENLAKAKEMLANSVKRE 367  
DB 139 KDHAQGLIGVTDENAKKAILANAGKDGVEELEKLSGSLSISKAKEMLANSVKRE 198  
QY 368 L 368  
DB 199 L 199  
RESULT 2  
US-09-974-992-7

```
Sequence 7, Application us/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel Ospc-derived peptide fragments
FILE REFERENCE: 459-666p
CURRENT APPLICATION NUMBER: us/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii
US-09-974-992-7
```

```
Query Match 32.7%; Score 588.5; DB 6; Length 212;
Best Local Similarity 68.8%; Pred. No. 2.7e-33;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNSGKDG-NASANSDESVGPNLFEISKRTESNAVLAVKEVELLASIDELATK 59
Db 17 ISCNNSGKGDSASTNPDESAGKPNLFEISKRTESNAVLAVKEVELLASIDELAK 76
QY 60 AIGKKT-GNNGLEANSKNTSLSGAVASDLIAEKLNVLNK-EELKEKIDTAKOCSTEF 117
Db 77 AIGOKIDNNNGIALNLNNGSILGAYASTLITERKLSKMLEKTEIAAKKCSSEF 136
QY 118 TTKKSEHAVLGLDNTDNOARAILIKKANKDKGALEKLFKAVENISKAADTLKNA 177
Db 137 TTKLSGSHADLKQDATTDHAKAAILKTHATTDKGAKKEFKLFESVEGILKAQVALTNS 196
QY 178 VKELTSPV 186
Db 197 VKELTSPV 205
```

```
RESULT 3
US-09-974-992-3
Sequence 3, Application us/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel Ospc-derived peptide fragments
FILE REFERENCE: 459-666p
CURRENT APPLICATION NUMBER: us/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii
US-09-974-992-3
```

```
Query Match 31.5%; Score 567; DB 6; Length 207;
Best Local Similarity 63.1%; Pred. No. 7.8e-32;
Matches 118; Conservative 33; Mismatches 32; Indels 4; Gaps 3;
```

```
QY 1 MACNSGKDGNASANSADSVKPNLFEISKRTESNAVLAVKEVELLASIDELATKA 60
Db 17 ISCNNSG--GDSASTNPDESAGKPNLFEISKRTESNAVLAVKEVELLASIDEL-SKA 73
QY 61 ICKTKIGNG-LEANSKNTSLSGAVASDLIAEKLNVLNK-EELKEKIDTAKOCSTEF 119
Db 74 ICKTKINDGTLEANKRNSLISLAGAVETSKRTQSLVNSEELKEKIDTAKOCSEKPT 133
QY 120 KIKSEHAVLGLDNTDNOARAILIKKANKDKGALEKLFKAVENISKAADTLKNAVK 179
Db 134 KIKDSHAEIGSQYDDNNAKKAAILKTHGTRKDGAKAELEFRSLSSLSKAQAALITNSVK 193
QY 180 ELTSPV 186
Db 194 ELTNPV 200

RESULT 4
US-09-815-242-5883
Sequence 5883, Application us/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: us/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5883
LENGTH: 837
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5883
```

```
Query Match 9.8%; Score 176; DB 6; Length 837;
Best Local Similarity 23.1%; Pred. No. 0.00026;
Matches 93; Conservative 77; Mismatches 154; Indels 78; Gaps 20;
```

```
QY 4 NNSGKDGNASANSADSVKPNLFEISKRTESNAV-VLAVKE-VELLASI----- 53
Db 100 NNLLEAVAKONANRAIDGLISLNGPQAKLKEOVGQATTLPNVTYVNDNAOTLVAKGIR 159
QY 54 DELATAIGKICGNNGLEANSKNTSLSGAVASDLIAE---KLNVKNEELKEKIDT 109
Db 160 DSIANEAT-IKAGQNTYDASQNKOTDYNSAVTAAKAIIIGQTSPSMNAQEIANDQV-T 217
QY 110 AKOCSEFTNKKL-----SEHAVLGLDNTDNOARAILIKKANKDKGALEKLFKAVE 164
Db 218 AKQOALNGEENLRTAQTNAKQHLNGLSLDLD--AOKDAVKROI---EGATHVNEYTOAN 272
```



Db 351 AVTQA-----EQLNKAQCPNPTSCKD---VETALENVQRAKELNGNQNANAK 396  
Qy 272 TISKLTOKLDGKNSER--LKEKIEENAK--CSEDFTKLGEHAQGIEN-VTDENA 325  
Db 397 TTKANALNLTGINNAQKALKSQIEGATTVAGVNOVSTSELMTANSLNQNGLNDEEA 456  
Qy 326 KKAILTDAKDKGALELEKLEKAVENLAKAKEMLANSVKE 367  
Db 457 TKAALNGTQNLKAKQHANTAIIDGLSHLTNQAQKALQVQ 498

## RESULT 7

US-10-072-851-13080  
; Sequence 13080, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C  
; FILE REFERENCE: ELITRA.028A  
; CURRENT FILING DATE: 2002-02-08  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 13080  
; LENGTH: 875  
; TYPE: PRP  
; ORGANISM: Staphylococcus aureus  
US-10-072-851-13080

Query Match 9.8%; Score 176; DB 7; Length 875;  
Best Local Similarity 23.1%; Pred. No. 0.00028;  
Matches 93; Conservative 77; Mismatches 154; Indels 78; Gaps 20;

Qy 4 NNSKRGASANSADSE--SVKGNLTIEISKITESNAV--VLAWE-VETLLASI----- 53  
Db 137 NNNLEVAQNANTAIIDGLTSLNGPQAKIKQVGAATLTPVQTVRDNAQTLLNTAMGLR 196  
Qy 54 DELATKAIGKIGNNGLEANSKNTSLSGAVATSDIAE---KLVAKMELEKEDYT 109  
Db 197 DSIAMEAT-ITAGONYDASQNKQTDYNSAVTAAKAIIGTTSQNSMAQETNOAKDY-T 254  
Qy 110 AKQOSTEFTNKLK-----SEHAVLGLDNLTDNQAORAILKKHANKDKGALELEKFAVE 164  
Db 255 AKQALNGQELRTAQTNAKQHLNGLSLDLD--AQKDAVKRQI---EGATVNEVTQAQN 309  
Qy 165 NLS--KKAQDTLKAQVLELSPYIVHGNNGKDGNTSANSADSESVKGNLTIEISKITESN 222  
Db 310 NADALNTAMNLTNGKIOD-----ONTIKQG--VNEFTDAE-----AKRNAVTN 350  
Qy 223 AVYLAWEIETILASIDELATKAIGKIIQONGGLAVEAG-----HNGTLGAY 271  
Db 351 AVTQA-----EQLNKAQCPNPTSCKD---VETALENVQRAKELNGNQNANAK 396  
Qy 272 TISKLTOKLDGKNSER--LKEKIEENAK--CSEDFTKLGEHAQGIEN-VTDENA 325  
Db 397 TTKANALNLTGINNAQKALKSQIEGATTVAGVNOVSTSELMTANSLNQNGLNDEEA 456

Qy 326 KKAILTDAKDKGALELEKLEKAVENLAKAKEMLANSVKE 367  
Db 457 TKAALNGTQNLKAKQHANTAIIDGLSHLTNQAQKALQVQ 498

## RESULT 8

US-09-708-427-15045  
; Sequence 15045, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. Alexandrov et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15045  
; LENGTH: 1304  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: 1..1304  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc-feature  
; LOCATION: 1..1304  
; OTHER INFORMATION: Ceres Seq. ID 1828628  
US-09-708-427-15045

Query Match 9.5%; Score 170; DB 6; Length 1304;  
Best Local Similarity 22.8%; Pred. No. 0.0012;  
Matches 91; Conservative 80; Mismatches 136; Indels 92; Gaps 17;

Qy 4 NNSKRGASANSADSE-----VKGNLTIEISKITESNAVYLAWE--EVEY 48  
Db 85 NEKLEKALAQKRAEESFEVEKFRAVELEQAGLEAVQKDVTSKNEDESIRSHALDISA 144  
Qy 49 LLASIDELATFAIGKIGNNGLEANSKNTSL--LSGAVATSDIAEKLKLVAKME----- 101  
Db 145 LLSITTEEL-----QRYKHEISMTADAKNNKLSHAEEYTKIAETHAEKAEILASLEGLRLK 198  
Qy 102 -----ELKEKIDTAKOSTEFTNKLKSEHAVLGLDNLTDNQAORAILKKHANKDKGALE 155  
Db 199 ALLGSKKEKALE-----GNEIVSKIKSE-----TELLRGLKELVSLIESLKKQEBGLVE 248  
Qy 156 LEKL-----FRAVENLSKRAQDTLKNAYKELTSPYVHGNNGKDGNTSANSADSEVKG--P 209  
Db 249 QLKVDLEAKKAESCNTSNVSVEWKNKYHELEKEV-----EESNRKSSASSESVMK 301  
Qy 210 NLTEISKITESNAVYLAWE--IETLLASIDELAT--KAIGKIIQONGGLAVEAGHNGTL 266  
Db 302 QLAELNVILHETKTDNAQAQKEKIELKRTIEAQTDLLEEGROY----- 345  
Qy 267 LAGAYTISKLTOKLDGKNSERLKEKIEENAKKSCSEDFTKLGEHAQGIEN-VTDENA 325  
Db 346 -----CIAKEBASKELEW--VESIKSELEISQ---EKTALDMEKAAVTSNIGQLDQRT 395  
Qy 326 KKAILTDAKDKGALELEKLEKAVENLAKAKEMLANSVKE 364  
Db 396 ELSTEL-----ERCKYVEERSKDMESLTLALOEASTES 429

## RESULT 9

US-09-708-427-15044  
; Sequence 15044, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. Alexandrov et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427



```

1 CURRENT FILING DATE: 2000-11-09
2
3 NUMBER OF SEQ ID NOS: 85364
4
5 SOFTWARE: PatentIn version 3.1
6
7 SEQ ID NO 15044
8
9 LENGTH: 1313
10
11 TYPE: prt
12
13 ORGANISM: Arabidopsis thaliana
14
15 FEATURE:
16
17 NAME/KEY: misc_feature
18
19 LOCATION: 1..1313
20
21 OTHER INFORMATION: Xaa is any amino acid
22
23 NAME/KEY: misc_feature
24
25 LOCATION: 1..1313
26
27 OTHER INFORMATION: Ceres Seq. ID 18286227
28
29 OS-09-708-427-15044

```

Query Match	9.5%	Score 170;	DB 6;	Length 1313;
Best Local Similarity	22.8%	Pred No. 0.0012;		
Matches 91;	Conservative 80;	Mismatches 136;	Indels 92;	Gaps 17

```

4 NNSGDSGMAANSNDSDE-----VKGNLTLEISKITTESNAUVLAVK-----EVET 48
Db 94 NEKLEKALAAKRADESEVEKEPRAVELEAGLEAVOKKDVTSTKNELESTIRSOHALDISA 153
OY 49 LIASIDELATAIKAKKIGNNGLLEANOSKNTSL--LSGAYASIDJIAKLVVLKNE----- 101
Db 154 LSTTEEL-----QRYVHELSTMDADKNNKALSHAEERTJAEIHAKKAEILASSELGRK 207
OY 102 -----ELKEKIDPAKOCSTFEFTKKLSEHAUVGLDNLUTDMMORAILKHNKDKGAE 155
Db 208 ALLSGEKEKEAIE-----GNEIYAKLKESE-----IELLRGLEKVSJLSESLKEOEGLVE 257
OY 156 LEKLT-----FRAVENLSKAAOULTNAYKELTSPVHGNNSGKOCNTSANSDESVMK--P 209
Db 258 QLKVDLEAKKAESOCNTSVEEMKNNKHLEKEY-----EESNNKSSSASMESVMK 310
OY 210 NLUTEISKITTESNAUVLAVKE--IETLVIASIDELAT--KAIGKAKIOQNGGLAVEAGHNGTL 266
Db 311 QLAELNHYHLHETKSDMAOQKKEIILELTETIEAQTDLDEBYGROY----- 354
OY 267 LAGAVTISKLTQKLDGLKNSSEKKEIENAKKCSDEFTKKLLEGHAOL--GIENVDEN 323
Db 355 -----CIAKEASKELENT--VESIKSELETSQ---EKKTRALDNEKAATSNIOULDOPT 404
OY 326 KKALITIDAADKGALELEKFAVENNAAKAKEMLANS 364
Db 405 ELSTIEL-----ERCKVEEERKSKRMESTJTLALQASSTES 438

```

```

: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,622
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows
: SEQ ID NO 5835
: LENGTH: 2434
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
: US-09-815-242-5835

```

Query Match	9.4%	Score	169.5	DB	6	Length	2434
Best Local Similarity	23.9%	Pred	No. 0.0029				
Matches	94	Conservative	76	Mismatches	144	Indels	79
						Gaps	21

```

QY      4 NNSKDSGNASANSADE----SVKPNNLTETSKRTESNAY--VLAVEK--VEFTLLASI-----53
Db      702 NHNIETAKONANNTALIDGUTSLNGPOKAKEGYCATTLTPNVQYTRDNAQTLLNTAMKGLR 761
QY      54 DELATRAIGKRRIKINNNGLEANOSKNSTSLSGAVAIISDLIAE----KLNYLKNKEELEKIDT 109
Db      762 DSIANEAT-IKAGONTVDASQNKOTDYNSAIVPAKAIIQOTTSPSMNAOETIOWADUY-T 819
QY      110 AKOCSTEFTKKLK-----SEHAVLGIDNLDTDDNAQAIIKKRKANKDKGALELEKIFRAVE 164
Db      820 AKOOALNGOEHLRTAQTNAAQHNLGSLDILT--AORDAVKROI-----EGATHNEVETOQN 874
QY      165 NLS--KAAQOTLNKVAVELTSPVHGNNSSGDONTSANGADESVKPNLTETSKRTESN 222
Db      875 NADPLMTAMNRLKNGIOD-----ONTIKOG-VNFTDADE-----AKRNATYN 915
QY      223 AVYLAVAEIEFTLLASIDELATRAIGKKIOONGSLAVEAG-----HNGTLLAGAY 271
Db      916 AVYQA-----EQILTKAQQPNTSHDG--VEFALENVORAKNELNOVVANAK 961
QY      272 TISLTLITOKLDGLKSNK--LKEKIEKNAK--CSEDFFPKLLGEHAQLGIEN-VTDENA 325
Db      962 TTAKNALNNLTISINNAQEKALKSQIGATTVAGVNVOYSTJASELNTAMSNLONGINDEEA 1021
QY      326 KKA-LITDAKDGALELEKFEKFAVENLAKKA 357
Db      1022 TKAAQVTTDADREKQTYNDNAVPAATLLDKTA 1054

```

%	TITLE OF INVENTION: Identification of Essential Genes In
%	TITLE OF INVENTION: Prokaryotes
%	FILE REFERENCE: ELITRA.011A
%	CURRENT APPLICATION NUMBER: US/09/815,242
%	CURRENT FILING DATE: 2001-03-21
%	PRIOR APPLICATION NUMBER: 60/191,078
%	PRIOR FILING DATE: 2000-03-21
%	PRIOR APPLICATION NUMBER: 60/206,848
%	PRIOR FILING DATE: 2000-05-23
%	PRIOR APPLICATION NUMBER: 60/207,727
%	PRIOR FILING DATE: 2000-05-26
%	
%	APPLICANT: Trawick, John D.
%	APPLICANT: Yamamoto, Robert T.
%	APPLICANT: Roemer, Terry
%	APPLICANT: Jiang, Bo
%	APPLICANT: Boone, Charles
%	APPLICANT: Bussey, Howard
%	TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
%	TITLE OF INVENTION: Proliferation
%	FILE REFERENCE: ELITRA.028A
%	CURRENT APPLICATION NUMBER: US/10/072,851
%	CURRENT FILING DATE: 2002-02-08



TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-10-072-851-12996

Query Match 9.4%; Score 169.5; DB 7; Length 6281;  
Best Local Similarity 23.9%; Pred. No. 0.0099;  
Matches 94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

OY 4 NNSGDKGNSANSAD---SVKPNLTETSKITESNAV--VLAYE-VETLLAST----- 53  
DB 1481 NHNLEKAKONANTYDGLTSLNGPOKAKLEQGVGATTPNVQVADNMQTLTAKGLR 1540  
OY 54 DELATFAIKKIKGNNGLENOSKNTSLSGAVASIDLI---KLVNLTNEELKEIKIT 109  
DB 1541 DSITANFAT-IKAGQNTYDASQNKQTDYNSAVTAAKAIIGQTPSPSNADOEINAKQOV-T 1598  
OY 110 AKQCSFTFNKLEK---SEHAVLGDNLTDDNAQRAIKKHNKDKGALEKLFKRAVE 164  
DB 1599 AKQALNGENLRTAOTNAKQHLNGSLDLD--AQKDAVKRQI---EGATHNEVYQAON 1653  
OY 165 NLS--KAAODTKNAVKELTSPYVHGNNSGDKGNTSANSADSVKGNLTETSKITESN 222  
DB 1654 NADALTMTNTKNGIOD-----QNTIKQG-VNFTDADE-----AKRNAYTN 1694  
OY 223 AVVLAVKEIETLLASIDELATKAIGKIKQONGGLAVEAG-----HNGTLLAGAY 271  
DB 1695 AVTQA-----EQILNKAQGNPNTSKD--VETALENVORAKNELNGONVANAK 1740  
OY 272 TISKLTQKLDGLKNSK--LKEIKENAKK---CSEDFTKLEGEHAQIGIEN-VTDENA 325  
DB 1741 TTKAKNALNLTSTNNAKQALKSQIGCATTVAGVNOVSTTASLNTAAMSLONGINDEAN 1800  
OY 326 KRAI-LITDAKDKGALEKLEKRAVENLAKAA 357  
DB 1801 TKAQKTYTDADREKOTYAVDVAATAKTLDDKTA 1833

RESULT 14  
US-10-037-182-8  
Sequence 8, Application US/10037182  
GENERAL INFORMATION:  
APPLICANT: Tryggvason, Karl  
APPLICANT: Dol, Masayuki  
TITLE OF INVENTION: Recombinant Laminin 10  
FILE REFERENCE: 99-274-F  
CURRENT APPLICATION NUMBER: US/10/037,182  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,449  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/279,282  
PRIOR FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1765  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-037-182-8

Query Match 9.4%; Score 168.5; DB 7; Length 1765;  
Best Local Similarity 24.6%; Pred. No. 0.0023;  
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

OY 2 ACNNSGKDC-NASANSADSVKGNLTET-----SKKITESNAVVLAVKREVEITLLAS 52  
DB 1384 SCSETGCGPNCTDEGERKCGPGCGGLVTVAHNAHQAMDLDQVLSLAVEEOLSKM 1443  
OY 53 IDELATKAIGKIKGNNGLENOSKNTSLSGAVASIDLIETLAEKLVNLTNEELKEIKITAKQ 112  
DB 1444 VSAKALRAD-----EAKQSAEDILLK-----TNATREKMD-KSNEELRLNLQIQRN 1488

OY 113 CSTEFTNKLKSEHAVLG-----LDNLITDDNAQRA-----ILKKNANKDG 152  
DB 1489 FLTQDSADLDSIEAVANEVLEKEMPESTPOQLQNLITDIREVESLSQVEYITLDHSAADTA 1548  
OY 153 AAELKLFKRAVENLSKAODT-----LKNAVELTSPYVHGNNSGDKGNTSANSADSV 206  
DB 1549 RAEM--LLEBAKRAKSATDVKTADMYKALEBAEKAAQVAAKAIK-----QADEDI 1599  
OY 207 KGNP--LTEISKITESNAVVL-AVKEIETLLASIDELATKAIGKIKQONGGLAVEAGHN 263  
DB 1600 QGTQNLITLTISETTAASEETLFNQSORISELERNEVELKRA-----AQNSG---BAEYI 1651  
OY 264 GTLLAGAYTISKL---ITQKLDGLKNSKLEKIEIEN--AKKCEDFTKLEGEHAQIGIE 318  
DB 1652 EKVY---YTVKQSAEDVKKTLDG-ELDEKVK-KVENLAKKTEESADARRAKAEMLQNEAK 1706  
OY 319 NVTDENAKKAILITDDAK-----DKGALEKLFKRAVENLAKKAEKMLA 362  
DB 1707 TLLAQANSKLQLLKDLERKEDNQRYLEDK-AQELARLEGEVSLKIDISQKVA 1759

RESULT 15  
US-09-938-275-6  
Sequence 6, Application US/09938275  
GENERAL INFORMATION:  
APPLICANT: Gerardo Castillo  
APPLICANT: Alan Snow  
TITLE OF INVENTION: Therapeutic and Diagnostic Applications  
TITLE OF INVENTION: The laminin and laminin-derived protein fragments  
FILE REFERENCE: PROTEO.P03  
CURRENT APPLICATION NUMBER: US/09/938,275  
CURRENT FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 1786  
TYPE: PRT  
ORGANISM: Homo Sapiens  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: Swissprot P07942  
DATABASE ENTRY DATE: 1988-08-01  
US-09-938-275-6

Query Match 9.4%; Score 168.5; DB 5; Length 1786;  
Best Local Similarity 24.6%; Pred. No. 0.0023;  
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

OY 2 ACNNSGKDC-NASANSADSVKGNLTET-----SKKITESNAVVLAVKREVEITLLAS 52  
DB 1405 SCSETGCGPNCTDEGERKCGPGCGGLVTVAHNAHQAMDLDQVLSLAVEEOLSKM 1464  
OY 53 IDELATKAIGKIKGNNGLENOSKNTSLSGAVASIDLIETLAEKLVNLTNEELKEIKITAKQ 112  
DB 1465 VSEAKLRAD-----EAKQSAEDILK-----TNATREKMD-KSNEELRLNLQIQRN 1509  
OY 113 CSTEFTNKLKSEHAVLG-----LDNLITDDNAQRA-----ILKKNANKDG 152  
DB 1510 FLTQDSADLDSIEAVANEVLEKEMPESTPOQLQNLITDIREVESLSQVEYITLDHSAADTA 1569  
OY 153 AAELKLFKRAVENLSKAODT-----LKNAVELTSPYVHGNNSGDKGNTSANSADSV 206  
DB 1570 RAEM--LLEBAKRAKSATDVKTADMYKALEBAEKAAQVAAKAIK-----QADEDI 1620  
OY 207 KGNP--LTEISKITESNAVVL-AVKEIETLLASIDELATKAIGKIKQONGGLAVEAGHN 263  
DB 1621 QGTQNLITLTISETTAASEETLFNQSORISELERNEVELKRA-----AQNSG---BAEYI 1672  
OY 264 GTLLAGAYTISKL---ITQKLDGLKNSKLEKIEIEN--AKKCEDFTKLEGEHAQIGIE 318  
DB 1673 EKVY---YTVKQSAEDVKKTLDG-ELDEKVK-KVENLAKKTEESADARRAKAEMLQNEAK 1727

Tue Mar 19 10:57:56 2002

us-09-596-746a-42.rapn

Page 8

```

Oy      319  NVTDENAKKALLITDAK-----DKGAEELEKFEKAVENAKAKAKEMLA 362
      :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db      1728 TLLAQANSKIQLDLDERKEYEDORYLEDR-AGELARLEGVEYRSLIKDISOKVA 1780

```

Search completed: March 18, 2002, 09:58:33  
Job time: 391 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:55 ; Search time 39.62 Seconds

(without alignments)  
340.552 Million cell updates/sec

Title: US-09-596-746a-42

Perfect score: 1798  
Sequence: 1 MACNNSGKGNNSANSANDES.....AVENLAKAKAKEMLANSVKEL 368

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	644.5	35.8	210	1	OSCI_BORBU
2	588.5	32.7	212	1	OSCI_BORBU
3	346	19.2	214	1	VM03_BORHE
4	325.5	18.1	215	1	VM03_BORHE
5	175	9.7	1790	1	US01_YEAST
6	173.5	9.6	1167	1	CAGA_HELPY
7	168.5	9.4	1786	1	LMBI_HUMAN
8	167.5	9.3	1013	1	AGOD_DROME
9	161.5	9.0	944	1	NUP1_YEAST
10	158.5	8.8	1433	1	REST_CHICK
11	158	8.8	1182	1	CGA2_HELPY
12	157.5	8.8	483	1	M6_STRPY
13	157.5	8.8	1251	1	RBP2_PLAYB
14	157	8.7	1391	1	MST2_DROME
15	155.5	8.6	2017	1	MYSN_DROME
16	155	8.6	2022	1	ANT1_ONCVO
17	154.5	8.6	2349	1	TPR_HUMAN
18	153	8.5	1786	1	LMBI_MOUSE
19	152	8.5	1639	1	LMG1_DROME
20	151.5	8.4	1427	1	REST_HUMAN
21	151.5	8.4	1966	1	MYSB_CAEBL
22	151.5	8.4	3210	1	CENF_HUMAN
23	150	8.3	1509	1	MYSN_ACACA
24	149.5	8.3	1978	1	MYHB_CHICK
25	148.5	8.3	564	1	M12_STRPY
26	148.5	8.3	679	1	YKR9_YEAST
27	148	8.2	492	1	M5_STRPY
28	147	8.2	1186	1	SMC_BACSU
29	146.5	8.1	1164	1	BAG_STRAG
30	146.5	8.1	1189	1	SCIT_CHICK
31	146.5	8.1	1230	1	SMC3_YEAST
32	146.5	8.1	1957	1	Y086_SCHPO
33	145.5	8.1	705	1	YNP9_CAEBL

34	145	8.1	662	1	TEPA_BACSU
35	145	8.1	1147	1	CGAL_HELPY
36	145	8.1	1186	1	CAGA_HELPY
37	143.5	8.0	539	1	M2A_STRPY
38	143.5	8.0	918	1	YMOB_CAEBL
39	143.5	8.0	1233	1	YF16_YEAST
40	143.5	8.0	1972	1	MYHB_HUMAN
41	143	8.0	724	1	HMHR_HUMAN
42	142.5	7.9	2116	1	MYS2_DICDI
43	142	7.9	1084	1	MYS2_RABIT
44	142	7.9	1526	1	MYS2_SCHPO
45	141.5	7.9	3110	1	LMA2_HUMAN

#### ALIGNMENTS

RESULT	ID	OSCI_BORBU	STANDARD	PRT	210 AA.
1	OSCI_BORBU	007337	15-DEC-1998 (Rel. 37, Created)		
AC	OSCI_BORBU	007337	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	OSCI_BORBU	007337	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OSCI_BORBU	007337	OUTER SURFACE PROTEIN C PRECURSOR (PC).		
GN	OSCI_BORBU	007337	OSPC OR BB19.		
OS	OSCI_BORBU	007337	Borrelia burgdorferi (Lyme disease spirochete).		
OC	OSCI_BORBU	007337	Plasmid Ip54.		
OX	OSCI_BORBU	007337	NCBI_TaxID-139;		
RN	OSCI_BORBU	007337	SEQUENCE FROM N.A.		
RP	OSCI_BORBU	007337	STRAIN-ATCC 35210 / B31;		
RC	OSCI_BORBU	007337	MEDLINE-9339332; PubMed-8478108;		
RX	OSCI_BORBU	007337	Wilske B., Praeger-Mursic V., Jauris S., Pradel I., Soutschek E., Schwab E., Wanner G.,		
RA	OSCI_BORBU	007337	*Immunological and molecular polymorphisms of OspC, an immunodominant		
RT	OSCI_BORBU	007337	major outer surface protein of Borrelia burgdorferi.;		
RL	OSCI_BORBU	007337	Infect. Immun. 61:2182-2191(1993).		
RN	OSCI_BORBU	007337	SEQUENCE FROM N.A.		
RP	OSCI_BORBU	007337	STRAIN-ATCC 35210 / B31;		
RC	OSCI_BORBU	007337	MEDLINE-94041630; PubMed-8225587;		
RX	OSCI_BORBU	007337	Padilla S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.,		
RA	OSCI_BORBU	007337	*Molecular characterization and expression of p23 (OspC) from a North		
RT	OSCI_BORBU	007337	American strain of Borrelia burgdorferi.;		
RL	OSCI_BORBU	007337	Infect. Immun. 61:5097-5105(1993).		
RN	OSCI_BORBU	007337	SEQUENCE FROM N.A.		
RP	OSCI_BORBU	007337	STRAIN-ATCC 35210 / B31;		
RC	OSCI_BORBU	007337	MEDLINE-96025162; PubMed-7494039;		
RX	OSCI_BORBU	007337	Fukunaga M., Hamase A.,		
RA	OSCI_BORBU	007337	*Outer surface protein C gene sequence analysis of Borrelia		
RT	OSCI_BORBU	007337	burgdorferi sensu lato isolates from Japan.;		
RL	OSCI_BORBU	007337	J. Clin. Microbiol. 33:2415-2420(1995).		
RN	OSCI_BORBU	007337	SEQUENCE FROM N.A.		
RP	OSCI_BORBU	007337	STRAIN-ATCC 35210 / B31;		
RC	OSCI_BORBU	007337	MEDLINE-98065943; PubMed-9403685;		
RX	OSCI_BORBU	007337	Frazer C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	OSCI_BORBU	007337	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	OSCI_BORBU	007337	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,		
RL	OSCI_BORBU	007337	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,		
RN	OSCI_BORBU	007337	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		

RA	utterback T., Matthey L., McDonald L., Artlach P., Bowman C.,	
RA	Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,	
RA	Smith H.O., Venter J.C.,	
RT	Genomic sequence of a Lyme disease spirochaete, Borrelia	
RT	burgdorferi.	
RL	Nature 390:580-586(1997).	
CC	-1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.	
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID	
CC	ANCHOR.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).	
CC	-----	
DR	EMBL; X69596; CAA49306.1; -	
DR	EMBL; 001894; AAA16058.1; -	
DR	EMBL; D49497; BAA08457.1; -	
DR	EMBL; AE000792; AAC66329.1; -	
DR	TIGR; BBH19; -	
DR	InterPro: IPR001800; Lipoprotein_6.	
DR	Pfam; PF01441; Lipoprotein_6; 1.	
DR	ProDom; PD001149; Lipoprotein_6; 1.	
DR	ProSite; PS00013; PROKAR_LIPOPROTEIN; 1.	
KV	Outer membrane; Lipoprotein; signal; Plasmid; Antigen;	
KV	Complete proteome.	
FT	SIGNAL	1
FT	CHAIN	18
FT	LIPID	19
FT	BY SIMILARITY.	
FT	OUTER SURFACE PROTEIN C.	
FT	N-ACETYL DIGLYCERIDE (BY SIMILARITY).	
SO	SEQUENCE	210 AA; 22340 MW; 7AFC93BF91777BF CRC64;

	Query Match	35.8%;	Score 644.5;	DB 1;	Length 210;
	Best Local Similarity	75.6%;	Fred. No. 5.3e-23;		
	Matches 136;	Conservative 15;	Mismatches 28;	Indels 1;	Gaps 1;
OY	189 NNSGSDGMTSASASAEVSKGPLTETISKRTESNNAVLAVERKIEFTLLASIDELATRAIKG	248			
Db	20 NNSGDGMTSASASADEVSKGPLTETISKRTIDSNAVLAVKEVALLSIDEIAKAIRKG	79			
OY	249 KIQQNGGLAVEAGHGCTTLACAYTTSKITQTQLDLDKNSEKKTEKIENAKCSEDEFTKL	308			
Db	80 KIHONNGDLENNHNGSLLAGAVAIYSTILTKQKLDLKN-EGLKEKIDAKKCSSEFTFKL	138			
C	309 EGEHAOLGIENVTDENAKKAILITDAANDKGAAELEKLFKAVENLAKAKEMLANSVKEL	368			
Db	139 KEKHIDLCKEGVTADADAKEAALIKTGNTKTGCAGEILGLFEESVEVLSSKAKEMLANSVKEL	198			
RESULT 2					
ID	OSC2_BORBU	STANDARD:	PRT;	212 AA.	
AC	008137;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete),				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia,				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	STRAIN=PRO:				
RX	MEDLINE=92219995; PubMed=1560779;				
RA	Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,				
RA	Soutschek E.;				
RT	"Molecular analysis and expression of a Borrelia burgdorferi gene				
RT	encoding a 22 kDa protein (pc) in Escherichia coli.";				
RL	Mol. Microbiol. 6:503-509(1992).				

RP	SEQUENCE OF 1-205 FROM N.A.	[2]
RC	STRAIN-DK26;	
RX	MEDLINE=94075528; PubMed=8253951;	
RA	Thaisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;	
RT	"Polymorphism in ospC gene of <i>Borrelia burgdorferi</i> and	
RT	immunoreactivity of OspC protein: implications for taxonomy and for	
RT	use of OspC protein as a diagnostic antigen.";	
RL	J. Clin. Microbiol. 31:2570-2576(1993).	
CC	-1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.	
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID	
CC	ANCHOR.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; X62162; CAA44093.1; -;	
DR	EMBL; X73624; CAA52003.1; -;	
DR	InterPro; IPR001800; Lipoprotein_6.	
DR	Pfam; PF00141; Lipoprotein_6; 1.	
DR	ProDom; PD001149; Lipoprotein_6; 1.	
KW	Outer membrane; lipoprotein; Signal; Plasmid; Antigen.	
FT	SIGNAL 1	
FT	CHAIN 18	
FT	BY SIMILARITY.	
FT	OUTER SURFACE PROTEIN C.	
FT	LIPID 19	
FT	19 N-ACYL DIGLYCERIDE (BY SIMILARITY).	
SO	SEQUENCE 212 AA; 22499 MW; C2066231PBF2E7D4 CRC64;	

```
Query Match      32.7% ; Score 588.5; DB.1 : Length 212;
Best Local Similarity    68.8%; Pred. No. 1.7e-20;
Matches 130; Conservative   20; Mismatches 36; Indels       3; Gaps          3;
```

OY    1 MACNNSGKDG-NASANSADESVKGPNLEISRKITETESNAVVAIVEVTLLSIDELATK 59  
     :::|||||: || | ||||| ||||| ||||| ||||| ||||| ::|||  
DB    17 ISCNNSGKGGSAGSNRPADSESAKNLFEISKITIDNSAFPLAKVEVTLVLSDELAKK 76

OY    60 AIGKKI-GNNGLEANQSKNTSLLSGAVALISDLIAEKLNVLYLN-EELKEKITPAQCSTEF 117  
     A IGGDINNNGLALINNONGSLLAGAVAIISTLTILEKTLNLNEELKEIKARAKCSEFF 136

OY    118 TNKLSEHNAVILGNLTPTDDNQRALTKHANRDKQAALLEKFRAVENLSKAADDTLKNA 177  
     ||||| || | : || |: ||| : ||| | ||| : ||::| ||| | || |:  
DB    137 TNRKSAGSDHGKODATPDTHAKAAALKTKHTATTDGAKEFKDFPVSVEGLLRAAOVALTNS 196

OY    178 VKELTSPIV 186  
     |||||:

DB    197 VKELTSPPV 205

RESULT      3  
VM24\_\_BORHE ID VM24\_\_BORHE STANDARD; PRT; 214 AA.

AC    P32778;  
DT    01-OCT-1993 (Rel. 27, Created)  
DT    01-OCT-1993 (Rel. 27, Last sequence update)  
DT    01-FEB-1994 (Rel. 28, last annotation update)

CN    DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.  
   VMP24.  
OS    Borrelia hermsii.  
OC    Bacteria; Spirochaetales; Spirochaetaceae; Borrella.  
OX    NCBI\_Taxid=140;  
RN    [1]  
RF    SEQUENCE FROM N.A.  
RC    STRAIN=SSP. HSI SEROTYPE 24;  
RX    MEDLINE=93133110; Pubmed=1484486;  
RA    Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
RT    "Subtelomeric expression regions of Borrellia hermsii linear plasmids"

[illegible]

```

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO YMP24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC DR EMBL; L04789; AAA22967.1; -
CC DR InterPro; IPR001800; Lipoprotein_6.
CC DR Pfam; PF01441; Lipoprotein_6; 1.
CC DR ProDom; PD001149; Lipoprotein_6; 1.
CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC KW Outer membrane; Lipoprotein; Signal; Plasmid.
CC FT SIGNAL 1 215 PROBABLE.
CC FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
CC SO SEQUENCE 215 AA; 23139 MM; 684C74D35F87C771 CnC64.
CC -----
Query Match 18.1%; Score 325.5; DB 1; Length 215;
Best Local Similarity 42.3%; Pred. No. 9.8e-09;
Matches 80; Conservative 32; Mismatches 62; Indels 15; Gaps 5;
QY 189 NNSGKDGNTSANSADSVKCP---NLTEISKKITESNAVVLAVKEITLLASIDELATK 244
Db 20 NNGGPE-----LKSDEVKASDGTVDLDAKISKIKIDASDFASVKEVHTLVKSIDELA-K 73
QY 245 AIGKKIQ-QONGGLAVAGHNGTLLAGATTISKLINQKIDGKN---SEKLEKRIENAKK 299
Db 74 AIGKKIKNDNSNFEDENDHNSGLAGVQVILTVYAKLTLSLEQIIGISDELKTEGVYAKK 133
QY 300 CSEDFKTKLEGGHGLGTEYVDENAKKAILITTDAAKDKGAELKFLKFAVENLAKAKE 359
Db 134 ESEAFVQVSKSHRDPLAKEGVTDAAKAILVTDSQTKKGAELKLTMTALDELLAKAND 193
QY 360 MLANSVKEL 368
Db 194 AVETVYKEL 202
-----
RESULT 5
USO1_YEAST STANDARD: PRT; 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN USO1 OR INT1 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.,
RT "A cytoskeleton-related gene, uso1, is required for intracellular
RT protein transport in Saccharomyces cerevisiae."
RL J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.

```

```

RN      [3]
RP      SEQUENCE OF 1-8 FROM N.A.
RL      Bai Y., Symington L.S.;
RA      Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC      COMPLEX.
CC      -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC      MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC      ER AND THE GOLGI COMPLEX.
CC      -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC      OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC      COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC      -1- SIMILARITY: BELONGS TO THE VDP/USOI/YBL047C FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X54378; CAA38253.1; -
DR      EMBL; L03188; AAB00143.1; -
DR      EMBL; U53668; AAB6659.1; -
DR      PIR; A38455; A38455.
DR      HSSP; P80220; 1DIP.
DR      SGD; S0002216; USOI.
DR      InterPro; IPR002017; Spectrin.
KW      Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT      DOMAIN 1 724
FT      DOMAIN 725 1790
FT      DOMAIN 465 487
FT      DOMAIN 991 1790
FT      DOMAIN 1172 1786
FT      CONFLICT 847 847
FT      CONFLICT 924 924
FT      CONFLICT 1253 1253
FT      CONFLICT 1319 1319
FT      CONFLICT 1461 1461
FT      CONFLICT 1581 1581
FT      CONFLICT 1600 1600
FT      CONFLICT 1661 1661
FT      CONFLICT 1772 1772
SQ      SEQUENCE 1790 AA; 206424 MW; 6CE2B216BFD4818 CRC64;

Query Match
Best Local Similarity 22.7%; Score 175; DB 1; Length 1790;
Matches 106; Conservative 66; Mismatches 158; Indels 136; Gaps 19;

OY      3 CNSGKDG-----NASANSDESVKPNLTETSKIT-----TESNAVYLVK 44
DB      929 CNIISKEHEHISKELVEYKSRFSDHNLV--AKLTEKLSLANNKYMQAENESLTKAVE 986
OY      45 E-----VETLASIDELATKAIGKIGNNGLAN--QSKNTSLSGAVYISDLIAEK 94
DB      987 ESKNESSIOLNLOKIDMSQEKENFOIERSIKNEIOLKKT-----ISLDEQK 1038
OY      95 LNVLKNEE-----LKEKIDTAKOCSTEFNKLKS-----EHAVALGLDNV- 133
DB      1039 EETISKSSKDEYESQISLKEKLETATANDENVNKKISLTKTREELAEALAYKKNLK 1098
OY      134 -----TDDNAORAT-----LKHANKDKG-----AAELEKIFKAVENLS 167
DB      1099 NELETETSEKALKVEENEHEHLKEEKIOLKEKATEFTKQOLNSLRANLESLEKHEDLA 1158
OY      168 ---RAAODTLKNAVVELSPIVHGNNSGKDGNTSANSADSVKGNP----- 210
DB      1159 AOKKYEIOIANKEKQYVEEISQLN-----DETTSQOENESIKKKNDLEGEVYKAMKSTS 1214
OY      211 LTESKTKTESNAVYLVAVEIE-----TLASIDELATKAIGKIQOQNGCLAVEAGH 262

```

```

DB      1215 EEOSNLKKSEIDALNIOIKELKKRNETNEASLSIESKVSSETV--KIKF---LQDECNF 1269
OY      263 NGTLLAGAYTISKIPIQKIDGLKNSKELKEKENAKKCSDFPKTKEGHAQIGIENVTD 322
DB      1270 K-----EKEVSLDEKIKLASDEKKNKYLELQKESKIKNEEDAKTLEKIQLEKI 1319
OY      323 ENAKKAILITDAAKDKGADELKFAVENILAKAKEMLANSVKEL 368
DB      1320 TNLK-----AKEKSELSRLKTKTSSEERKNAEOLERKNEI 1358

RESULT 6
CAGA_HELPUP
ID      CAGA_HELPUP STANDARD; PRT; 1167 AA.
AC      GQZLTI;
DT      20-AUG-2001 (Rel. 40, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)
DE      (CAG PATHOGENICITY ISLAND PROTEIN 26).
GN      CAGA OR CAI OR CAG26 OR JHP0495.
OS      Helicobacter pylori J99 (Campylobacter pylori J99).
OC      Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC      Helicobacter.
OX      NCBI_TaxID=85963;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99120557; PubMed=9923682;
RA      Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA      Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA      Tummino P.J., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,
RA      Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA      Trust T.J.;
RT      "Genomic sequence comparison of two unrelated isolates of the human
RT      gastric pathogen Helicobacter pylori."
RL      Nature 357:176-180(1999).
CC      -1- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC      OR FUNCTION OF THE CYTOTOXIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE001483; AAD06073.1; -
DR      Antigen; Complete proteome.
KW      Antigen; Complete proteome.
FT      DOMAIN 246 249
FT      DOMAIN 882 889
FT      DOMAIN POLY-ASN.
SQ      SEQUENCE 1167 AA; 129729 MW; FD5E86B81CEBD0F2 CRC64;

Query Match
Best Local Similarity 9.6%; Score 173.5; DB 1; Length 1167;
Matches 105; Conservative 81; Mismatches 140; Indels 147; Gaps 25;

OY      5 NSGKGNASANSADSVKPNLTETSKITTESNAVYLVK-----VETLASID 54
DB      693 NINMDLKFSSPDEFKNGKN-KDPSKAPEFLKALKGSVKDGINPEWISVENYENMAALN 751
OY      752 EF-----KNGKKNDFSKVTQ-----AKSDLENSIKVDIINOKITDKVDNLNAV 795
DB      113 -----CSTFTNKLKSEHAYVLGIDNITDQN--QORALIKKHAKKDKGADELKFAVEN-- 165
DB      796 SVAKATGDFSG---VEQALADLKNSKSKQOLAQAQAKNDPFGKNKA---LYOSVKNQV 848
OY      166 -----LSRAAODTL-----KNAVVELSPIVHGNNSGKDGNTSANSADSVKGNP 213
DB      849 NCTLVGNLSRAEATITLSKNFSDIKELNAKL--GNFNNNNNGLENSTE-----PIYIQ 901

```



```

OY 214 ISITSESNVVLAVLVEIPLLASIDELTPKAIGK--KIQQNGGLAVEAGNH--GTL-- 266
Db 902 VAKV---KAKIRLDQIASGLSDGQAAASFLLKRRDKVDLSKGLSANHEPIYATID 958
OY 267 LAGAVYTSK-----LITQKLDGLKNS-----EKLKER- 293
Db 959 LGGFPLKRRDKVDLSKGLSREKLTQKIDNLNQAASEAKASHFDNLDMIDKLKNS 1018
OY 294 -----IENAKKCEDFPFKLE---GEHQQL--GIEN-----VTDENAKKAI 329
Db 1019 KKNVNVLYVESAKKVPFPTSLSAKIDNATNTSHRINSNVKNGTINEKAGMGLTKNSEWLK 1078
OY 330 LITD--AAKKGAAELEKLFKAVENTLAKAKACM-----LANSVEL 368
Db 1079 LVNDKIYAAHVGSAPL-----SAYDKTGFGNKKMKKDYSDSFSTKPLSNVADKI 1127

RESULT 7
LMB1_HUMAN STANDARD; PRT; 1786 AA.
P07942;
01-AUG-1988 (Rel. 08, Created)
01-AUG-1988 (Rel. 08, Last sequence update)
20-AUG-2001 (Rel. 00, Last annotation update)
DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN LMB1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP MEDLINE=90368768; PubMed=1975589;
RX MEDLINE=90368768; PubMed=1975589;
RA Vuolteenaho R., Chow L.T., Tryggvason K.;
RT "Structure of the human laminin B1 chain gene.";
RL J. Biol. Chem. 265:15611-15616(1990).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=87280097; PubMed=3611077;
RA Pikkatainen T., Eddy R., Fukuishima Y., Byers M., Shows T.,
RA PhilpottJenlett T., Saraste M., Tryggvason K.;
RT "Human laminin B1 chain. A multidomain protein with gene (LMB1)
RT locus in the q22 region of chromosome 7.";
RL J. Biol. Chem. 262:10454-10462(1987).
RN 13
RP SEQUENCE OF 1276-1709 FROM N.A.
RX MEDLINE=88021029; PubMed=3661559;
RA Jaffe M., Modl W.S., Ricca G.A., Mudd R., Chiu I.M., O'Brien S.J.,
RA Drosan W.N.;
RT "Isolation of a cDNA clone for the human laminin-B1 chain and its
RT gene localization.";
RL Am. J. Hum. Genet. 41:605-615(1987).
RN 14
RP FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ
RN IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
RN CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
RN WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
RN 15
RP SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
RN DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
RN TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
RN COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
RN THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-2
RN (MEROSIN), AND LAMININ-6 (K-LAMININ).
RN 16
RP SUBCELLULAR LOCATION: EXTRACELLULAR.
RN 17
RP TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
RN COMPONENT).
RN 18
RP DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
RN WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
RN 19
RP DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
RN 20
RP SIMILARITY: CONTAINS 1 AND IV ARE GLOBULAR.
RN 21
RP SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
RN 22
RP SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.

```

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL, M61951, AAAS9486.1, JOINED.
DR	EMBL, M58147, AAAS9486.1, JOINED.
DR	EMBL, M61917, AAAS9486.1, JOINED.
DR	EMBL, M61918, AAAS9486.1, JOINED.
DR	EMBL, M61921, AAAS9486.1, JOINED.
DR	EMBL, M61922, AAAS9486.1, JOINED.
DR	EMBL, M61923, AAAS9486.1, JOINED.
DR	EMBL, M61924, AAAS9486.1, JOINED.
DR	EMBL, M61925, AAAS9486.1, JOINED.
DR	EMBL, M61926, AAAS9486.1, JOINED.
DR	EMBL, M61927, AAAS9486.1, JOINED.
DR	EMBL, M61928, AAAS9486.1, JOINED.
DR	EMBL, M61929, AAAS9486.1, JOINED.
DR	EMBL, M61930, AAAS9486.1, JOINED.
DR	EMBL, M61931, AAAS9486.1, JOINED.
DR	EMBL, M61932, AAAS9486.1, JOINED.
DR	EMBL, M61933, AAAS9486.1, JOINED.
DR	EMBL, M61934, AAAS9486.1, JOINED.
DR	EMBL, M61935, AAAS9486.1, JOINED.
DR	EMBL, M61936, AAAS9486.1, JOINED.
DR	EMBL, M61937, AAAS9486.1, JOINED.
DR	EMBL, M61938, AAAS9486.1, JOINED.
DR	EMBL, M61939, AAAS9486.1, JOINED.
DR	EMBL, M61940, AAAS9486.1, JOINED.
DR	EMBL, M61941, AAAS9486.1, JOINED.
DR	EMBL, M61942, AAAS9486.1, JOINED.
DR	EMBL, M61943, AAAS9486.1, JOINED.
DR	EMBL, M61944, AAAS9486.1, JOINED.
DR	EMBL, M61945, AAAS9486.1, JOINED.
DR	EMBL, M61946, AAAS9486.1, JOINED.
DR	EMBL, M61947, AAAS9486.1, JOINED.
DR	EMBL, M61948, AAAS9486.1, JOINED.
DR	EMBL, M61949, AAAS9486.1, JOINED.
DR	EMBL, M61950, AAAS9486.1, JOINED.
DR	EMBL, M55370, AAAS9485.1, JOINED.
DR	EMBL, M55378, AAAS9485.1, JOINED.
DR	EMBL, M53365, AAAS9485.1, JOINED.
DR	EMBL, M55371, AAAS9485.1, JOINED.
DR	EMBL, M55372, AAAS9485.1, JOINED.
DR	EMBL, M55373, AAAS9485.1, JOINED.
DR	EMBL, M55374, AAAS9485.1, JOINED.
DR	EMBL, M55375, AAAS9485.1, JOINED.
DR	EMBL, M55376, AAAS9485.1, JOINED.
DR	EMBL, M55344, AAAS9485.1, JOINED.
DR	EMBL, M55345, AAAS9485.1, JOINED.
DR	EMBL, M55346, AAAS9485.1, JOINED.
DR	EMBL, M55347, AAAS9485.1, JOINED.
DR	EMBL, M53348, AAAS9485.1, JOINED.
DR	EMBL, M53349, AAAS9485.1, JOINED.
DR	EMBL, M55350, AAAS9485.1, JOINED.
DR	EMBL, M55351, AAAS9485.1, JOINED.
DR	EMBL, M55352, AAAS9485.1, JOINED.
DR	EMBL, M55353, AAAS9485.1, JOINED.
DR	EMBL, M53355, AAAS9485.1, JOINED.
DR	EMBL, M53356, AAAS9485.1, JOINED.
DR	EMBL, M53357, AAAS9485.1, JOINED.
DR	EMBL, M55358, AAAS9485.1, JOINED.
DR	EMBL, M55359, AAAS9485.1, JOINED.
DR	EMBL, M55360, AAAS9485.1, JOINED.
DR	EMBL, M53361, AAAS9485.1, JOINED.
DR	EMBL, M53362, AAAS9485.1, JOINED.
DR	EMBL, M55363, AAAS9485.1, JOINED.
DR	EMBL, M55364, AAAS9485.1, JOINED.
DR	EMBL, M55365, AAAS9485.1, JOINED.
DR	EMBL, M53367, AAAS9485.1, JOINED.
DR	EMBL, M53368, AAAS9485.1, JOINED.

FT	DISULFID	886	895	BY SIMILARITY.
Query Match			9.4%	Score 168.5; DB 1; Length 1786;
Best Local Similarity			24.6%	Pred. No. 0.99;
Matches 102;	Conservative	61;	Mismatches 160;	Indels 91; Gaps 21;
OY	2	ACNNSKGDG-NASANASDESXKGPNUDEL-----SKKITESNAVYLAKVEYETLLAS	52	
Db	1405	SCSETECGGPNCHTDEGKERCGPGCCGLVYVAHNAWQKRAMDLDDVYSLALAEVQSLSKM	1464	
OY	53	IDELATKAIGKIGKIGNGLEANOSKNTSLSGAYASIDLIAKLNYLKNKELEKEDTFAKQ	112	
Db	1465	VSEAKTRRAD-----EAKOSADIEDLLK-----TNAFTKEKMD-KSNEELRNLIKQIRN	1509	
OY	113	CSETEFNKTKSEHAVLVC-----LDNLTDNAQQA-----ILKIRANKDG	152	
Db	1510	FLTQDSADLDSTEAVANLEYLKNEMPSTPOQJONTEDTREVESLSOVEYTLQHSADIA	1569	
OY	153	AAELEKLFKAVENLSKAQDT-----LKNVKELTSPYVGNNSGKDGNTSANSASDV	206	
Db	1570	RAEM-LLEAKRASKSATDYKVTADMYKKALEAEKKAQVAEAKIR-----QADEDI	1620	
OY	207	KAPN-LTDEIKSKITEESNAVYL-AKVEETTLASIDELATKAIGKIQONGSLAVEAGHN	263	
Db	1621	QSTQMLNLTISEETVASEETLFEMASQRISELERVVEELKRRK-----AQNNG---EAEYL	1672	
OY	264	GTLLAGAVTYISKL---ITQKLDGLKNSSEKTEKTEEN-AKKSCEDPFKTEGGAQGLTE	318	
Db	1673	EKVY---TYVQSADEVKKTLIDG-ELDEKYK-KVENLIAKTEESADARRAKAEMLQNEAK	1727	
OY	319	NVTDENAKKALIIDPAK-----DKGALEKLFKAVENLAKAKEMLA	362	
Db	1728	TLLAANSKIQTLNDLEKRYEDNQRYLEDK-AQELTARLEGEVSRLIKDISQKVA	1780	
RESULT	8			
AD6D_DROME				
ID	AD6D_DROME	STANDARD;	PRT;	1013 AA.
AC	P91927; Q9W160;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CALCIUM-BINDING MITOCHONDRIAL PROTEIN ANON-60DA.			
GN	ANON-60DA OR CG4589.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Phylogota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
NCBI	TaxID=7227;			
OX	[1]			
RP	SEQUENCE FROM N.Y.			
RC	STRAIN=BRKRELEY.			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Hamann-Lindes P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson S.C.R., Miklos G.L.G.,			
RA	April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,			
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Burtis J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Mayan-Hocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Giolek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Iidegawa C.,			
RA	Talbot M., Kalnuch E., Klenov P., Klenov P., Klenov P.,			

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulo G., Milstine N.V., Modarity C., Morris J., Mostreli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamas I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkies R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RP [2]  
 RN SEQUENCE OF 626-944 FROM N.A.  
 RC TISSUE-Ovary;  
 RC MEDLINE=99168769; PubMed=10071211;  
 RA Caggese C., Ragone G., Perrini B., Moschetti R., de Pinto V.,  
 RA Calzai R., Barsanti P.;  
 RT "Identification of nuclear genes encoding mitochondrial proteins:  
 RT isolation of a collection of *D. melanogaster* cDNAs homologous to  
 RT sequences in the Human Gene Index database.";  
 RL Mol. Gen. Genet. 261:64-70(1999).  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO  
 CC FRAMESHIFTS IN POSITIONS 920 AND 930.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AE003464; AAF47217.1; -  
 DR EMBL: Y10912; CA71853.1; ALT\_FRAME.  
 DR Flybase: FBgn0019886; CG4598.  
 DR InterPro: IPR001813; 60s\_ribosomal.  
 DR InterPro: IPR002024; Bacterioferritin.  
 DR InterPro: IPR003871; DUF223.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR000861; REM\_repeat.  
 DR InterPro: IPR002555; RFL.  
 DR InterPro: IPR001236; 1dh.  
 DR Pfam: PF00428; 60s\_ribosomal; 1.  
 DR Pfam: PF02721; DUF223; 1.  
 DR Pfam: PF00036; ehand; 2.  
 DR Pfam: PF02185; HRI; 1.  
 DR Pfam: PF00056; 1dh; 1.  
 DR Pfam: PF01605; RFL; 1.  
 DR SMART: SMO0054; Eph; 2.  
 DR PROSITE: PS00018; EF\_HAND; 2.  
 KM Mitochondrion; Calcium-binding; Repeat.  
 FT CA\_BIND 700 711 EF\_HAND 1 (POTENTIAL).  
 FT CA\_BIND 783 794 EF\_HAND 2 (POTENTIAL).  
 FT CONFLICT 687 687 E -> Q (IN REF. 2).  
 FT CONFLICT 690 690 K -> G (IN REF. 2).  
 FT CONFLICT 717 717 K -> H (IN REF. 2).  
 FT CONFLICT 736 736 E -> D (IN REF. 2).  
 FT CONFLICT 740 740 K -> R (IN REF. 2).  
 FT CONFLICT 761 761 K -> E (IN REF. 2).  
 FT CONFLICT 943 943 G -> S (IN REF. 2).  
 SQ SEQUENCE 1013 AA; 113579 MW; 8C710380263F262E CRC64;

Query Match 9.3%; Score 167.5; DB 1; Length 1013;  
 Best Local Similarity 22.4%; Pred. No. 0.6; Mismatches 147; Indels 57; Gaps 11;  
 Matches 77; Conservative 63; Mismatches 147; Indels 57; Gaps 11;  
 QY 29 ISKRTESNAVLAKEVEITLASIDELATRAIKKIGNGNGLEANSQKNTSLSGAVALS 88  
 DB 609 VKPEVRESRAAKLLYNVNMMISGLDVL-----NDLEARQHQINQAESDVAAS 658  
 QY 89 DLIAEKLVKNEBLEKIDTAKOCSTEFNNKLSENAVJGLDNDNARAILKRAAN 148  
 DB 659 SPTEVPOAMVHIDELVATIRMEKASDEPERFVGGD-----LVKLEAD 702  
 QY 149 KDGALELEKLFRAVENLSAADDTLKNAYKELT-----SPIYGNNGSGDN 196  
 DB 703 KD-GVYSVNETRAVOSIDREATNIDKKQLEETELLASKASRRHEIYHIDLMNNIK 761  
 QY 197 TSANSADESQVGNLIEISKRTESNAVLAKEVEITLASIDELATRAIKKIQONGGL 256  
 DB 762 VLKETSDSEA-RLKHIEAVLEKFDADKGVYTVNDIRRVLESIGRDNIKLSDAIEE---- 816  
 QY 257 AVEAHNGHTLLAGAVTTSKLITOKLDLKNSEKIEENAKKSEDPFKLESEH-AQL 315  
 DB 817 LISLDEKQVLAQAEQKIEKAIAK---SMKEAEKLEKSEVDKADK---DLSKLVNDIHDSAK 870  
 QY 316 GIENVTDENAKKAILITDAKDKGALELEKLFRAVENLAKRAKE 359  
 DB 871 EIIDIANEMRDKERTVDPKAKELKA---EPAFK---DTAKTLKD 908  
 RESULT 9  
 ID NUFL\_YEAST STANDARD; PRT; 944 AA.  
 AC P32380;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE NUFL PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).  
 GN NUFL OR SPC110 OR YDR356W OR D9476.3.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C;  
 RA MEDLINE=92176232; PubMed=1541631;  
 RA Mirzayan C., Copeland C.S., Snyder M.;  
 RT "The NUFL gene encodes an essential coiled-coil related protein that  
 RT is a potential component of the yeast nucleoskeleton.";  
 RL J. Cell Biol. 116:1319-1332(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94064779; PubMed=7503995;  
 RA Kilmartin J.V., Dyos S.L., Kershaw D., Finch J.T.;  
 RT "A spacer protein in the Saccharomyces cerevisiae spindle poly body  
 RT whose transcript is cell cycle-regulated.";  
 RL J. Cell Biol. 123:1175-1184(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / AB972;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
 RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,  
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,  
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,  
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,  
 RA Rifken L., Riles L., Taich A., Trevisan E., Vignati D.,  
 RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL  
 CC ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT  
 CC IS ESSENTIAL FOR GROWTH.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE  
 CC NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE

CC NUCLEOLUS.  
 CC -1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: 211582; CAAT7668.1; -  
 DR EMBL: X73287; CAAS1733.1; -  
 DR EMBL: U28372; AAB64791.1; -  
 DR PIR: S26710; S26710.  
 DR PIR: S34288; S34288.  
 DR SGD: S0002764; NUP1.  
 DR Colled coll: Nuclear protein; Phosphorylation.  
 DR DOMAIN 164 791 COILED COIL.  
 DR DOMAIN 54 59 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 DR DOMAIN 726 731 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 DR DOMAIN 742 747 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 DR DOMAIN 731 944 ARG/TYR-RICH.  
 SQ SEQUENCE 944 AA; 111781 MW; 04FAA074BB8A0BC8 CRC64;

Query Match 9.0%; Score 161.5; DB 1; Length 944;  
 Best Local Similarity 23.2%; Pred. No. 1;  
 Matches 100; Conservative 71; Mismatches 167; Indels 93; Gaps 21;

DB 6 SCKDNASANSDESVKGNLFEISKITTESNAV-VLAVKEVETLASTIDE-LAKKAIGK 63  
 DB 91 SRRSGN-VDSKRRLIDDLKKDVPMSOPLEQREHOMKERRDRLESKLLGK 145  
 QY 64 K-----IGNGLEANSKNTSLSGAVASIDLAELKLVNKN-EELKEKIDYAKCS 114  
 DB 146 RHITVANSDISKELYINEIKSLK-----HEIKELREKENDTLNVDLEETDCLKNRL 200  
 QY 115 TETFKLSEHNAVGLDNLTDNAORALIKKHANDKAAELEKFAV-----EMLS 167  
 DB 201 QALEKELDAKKNIVN-SRRVDDHS--GCIEEOMERKIALELERKLVKDOVLELNN 257  
 QY 168 -----KAAODTLKNAVEKITSPVHGNNSKDGNTS-----ANSAD-SVKGP 210  
 DB 258 DVQSLKLSKDEKLKMLNLENE--LKSNAEKOTOLEFKKNELKRNELNELKIKSDE 315  
 DB 211 L-TETSKKITESNAVLAIVKKEITL-----ASIDELATRAIGKKIQO-----NGC 255  
 DB 316 MDLQKOKONESKRLKDELNELETFESENSGOSAKENELKMLKIALEELKEESTKNSQ 375  
 QY 256 LAVEAGHNGTLLAGAVTTSKITLTKDGL-----KNSKLEKIKENAKKSE----- 302  
 DB 376 LIAREKGLASLAPOULTQESKLNDRDQSLRSREELKTKTNKLDRIAAEEIVSKDER 435  
 QY 303 -DFTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALE-E--KLFAVENILAKAA 357  
 DB 436 IIDQKKVKQLENDLFVYKTHSEK---TITD-----NELSKOKLILILENDLKVA 485  
 QY 358 KEMLANSVKEL 368  
 DB 486 QEKYSKMEKEL 496

RESULT 10  
 REST\_CHICK STANDARD; PRT; 1433 AA.  
 AC 042184; 042228; 057563; 057564;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).  
 GN RSN.  
 OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus.  
 CC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98137792; PubMed=9469933;  
 RA Griparic L., Volosky J.M., Keller T.C. III;  
 RT "Cloning and expression of chicken CLIP-170 and restin isoforms."  
 RN Gene 206:195-208(1998).  
 RN [2]  
 RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).  
 RC TISSUE=Pectoralis muscle;  
 RA Griparic L., Keller T.C. III;  
 RT "Identification and expression of two novel CLIP-170/Restin isoforms  
 RT expressed predominantly in muscle."  
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN  
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE  
 CC CYTOSKELETON (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF014012; AAC60344.1; -  
 DR EMBL: AF020764; AAC60345.1; -  
 DR EMBL: AF045650; AAC03547.1; -  
 DR EMBL: AF045651; AAC03548.1; -  
 DR InterPro: IPR000938; CAP-GLY.  
 DR InterPro: IPR001878; ZnF\_CCHC.  
 DR Pfam: PF01302; CAP-GLY; 2.  
 DR SMART: SM00343; ZnF\_C2HC; 1.  
 DR PROSITE: PS00845; CAP-GLY 1; 2.  
 KW Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.  
 FT DOMAIN 79 121  
 FT DOMAIN 144 207  
 FT DOMAIN 235 277  
 FT DOMAIN 305 332  
 FT DOMAIN 351 1353  
 FT DOMAIN 1414 1427  
 FT VARSPPLIC 458 492  
 FT VARSPPLIC 458 492  
 FT VARSPPLIC 803 803  
 FT VARSPPLIC 458 458  
 FT CONFLICT 309 309  
 FT CONFLICT 440 440  
 FT SEQUENCE 1433 AA; 161026 MW; 5631CE8683498E23 CRC64;

Query Match 8.8%; Score 158.5; DB 1; Length 1433;  
 Best Local Similarity 22.5%; Pred. No. 2,2; Indels 67; Gaps 16;  
 Matches 86; Conservative 78; Mismatches 152;

QY 18 DESVKGPNLFEISKITTESNAVLAIVKKEITLASEIDELATKAIGKIGNGGLEANSKN 77  
 DB 508 DLALRVAEVALNGRLESSK-----HIDVDVTSLSLQETIS--SLQEMAAAGKE-HQREM 560  
 QY 78 TSLISGAVASIDLAELKLVN-----KNELEKIDYAKOCSTFTN--KLSEHA 126  
 DB 561 SLIKEKESSEALRKIKTLASNERMGKENSLETKLDHANKENSVDVLEKRSLESA 620  
 QY 127 VIGLNDNLTDNAORALIKKHANDKGA--AELEKFAVENLSKAODTLKN----- 176

```

Db 621 IASHQANEE-----LKVSPFKGVGAQTAEFAELKTQMEKVKLDYENMSMLTKQENE 674
Qy 177 -----AVKELTSPVHGNNSGKDGNTSANSADSVKGNLTETSKKTTESNAVLAWE 230
Db 675 KSOHLKEIEALKAKLLETVEEKEOTLEMLKAKLESVEDQHLVEMEDTLKLOEAEIKVE 734
Qy 231 IETLASIDELATKAIGKIKIOONGLAWEAGHNGTL-----LAGATYISKLITOKL-DGLK 285
Db 735 LDVLOAKNE-OTKLIGSLTQO-----IRASEEKLIDLALOKANSEGLKLEOKLSEQLQ 788
Qy 286 NSEIKERIEAKKSEDEFTKLEGEHAQIGENTDENAKKAILITDAKDKGAEL-- 343
Db 789 AAEKIOULETEK--VSNLTRELOKREOKL-----LDLEKNLSAVNGVKRSLKEKELQ 839
Qy 344 --EKLEKAV--ENLAKAKEML 361
Db 840 LKEKFTSAVDGAENQRAMQETI 862

```

```

RESULT 11
CGA2_HELPY STANDARD: PRT: 1182 AA.
P55746:

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN).
CN CACA OR CAI.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 53726 / 84-183;
RX MEDLINE-93239281; PubMed-8478069;
RA Tumuru M.K.R., Cover T.L., Blaser M.J.;
RT "Cloning and expression of a high-molecular-mass major antigen of
RL Helicobacter pylori: evidence of linkage to cytotoxin production.";

```

```

-1- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
OR FUNCTION OF THE CYTOTOXIN.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; L11714; ?; NOT_ANNOTATED_CDS.

```

```

DR HSSP; P02937; IMLP.

```

```

KW Antigen.
FT DOMAIN 878 885 POLY-ASN.
SO SEQUENCE 1182 AA; 131503 MW; C916817E2E57B84 CRC64;

```

```

Query Match 8.8%; Score 158; DB 1; Length 1182;
Best Local Similarity 23.8%; Pred. No. 1.9; Indels 122; Gaps 20;
Matches 101; Conservative 61; Mismatches 140;

```

```

Qy 24 PNLTEIS-----KKITESNAVLAWEVETLLASIDELATKAIGKIKNGGLEA 72
Db 541 PNLNLATLSVVRDLLEKLIKGLSPOEANKLVDFLSSNNKEIVKAL-----NENKAV 595
Qy 73 NOSKATSLISGAY-----AISDLIAEKLVLKNEBELKEKIDTAQOCSTFTYKLSSEHAV 127
Db 596 AEAKNT-----GNYDEVKRAQKDL--EK--SLKREHLEKGDVAKNLESISGKNKKEAR- 646
Qy 128 LGLDMLTDOMNOR-----AIIKKHANKD-----KG-----AAELKELKRAVENLSK 168

```

```

Db 647 -----AQANSQKDEIFALINKEANDARAIAVQNLKGIKRELSDKLENINKDLDFSK 700
Qy 169 A-----AQDT-----LKNAYEL-----TSPVHGNNSGKDG 195
Db 701 SEDGFKNRKNDFSKAEFLKALKGSVKDGINPEWISKEYENLNAALNEFKNKNDFSK 760
Qy 196 NTSANSADSVKGNLTETSKKTTESNAVLAWE-----EIEPLASIDELATKAIG 247
Db 761 VTQANSDOENSTIYVITNOKITDKYDELNOAVSAKICDFSGEALDLKLFKSEQLA 820
Qy 248 KRIIOONGL-----AVEAGHNGTLTAGATYISKLITOKLDGLKNSKLEKIEENAK 298
Db 821 QQAQKNESFNKCKSELYQGVKNNGVNTLVNG-----LSGIEATYALANFSDIK 870
Qy 299 KCSDEFTKLEGEHAQIGENTDENAKKAILITDAKDKGAELKFAVENL 353
Db 871 ELNEKF--KNFNNNNNNGGLKNGEPIYAQVNNKKKTGOVASSPEPIYAQVAKKVTKKIDOL 928
Qy 354 AKAA 357
Db 929 NQAA 932

```

```

RESULT 12
M6_STRPY STANDARD: PRT: 483 AA.
P08089;

```

```

DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 6 PRECURSOR.
GN EMB6.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13114;

```

```

RP SEQUENCE FROM N.A.
RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
RX MEDLINE-86111835; PubMed-3511046;
RT "Complete nucleotide sequence of type 6 M protein of the group A
RL Streptococcus. Repetitive structure and membrane anchor.";
J. Biol. Chem. 261:1677-1686(1986).

```

```

[2]
SEQUENCE OF 43-122 FROM N.A.

```

```

RX MEDLINE-85166224; PubMed-3885219;
RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;

```

```

RT "Relationship of M protein genes in group A streptococci.";

```

```

-1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
PHAGOCYTOSIS.

```

```

-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.

```

```

-1- SIMILARITY: TO OTHER M PROTEINS.
-1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
IN THE REGION OF THE MEMBRANE ANCHOR.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; M11338; AAA26920.1; -
DR PIR; A26297; A26297.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M; 9.
DR PRINTS; PR00015; GPOSANCHOR.

```

DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; 1.  
 KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;  
 KW Transmembrane; Coiled coil; Signal.  
 FT SIGNAL 1 42  
 FT CHAIN 43 483 M PROTEIN, SEROTYPE 6.  
 FT DOMAIN 43 457 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 458 457 MEMBRANE ANCHOR.  
 FT DOMAIN 478 477 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 69 138 10 x 7 AA TANDDEM REPEATS.  
 FT DOMAIN 157 269 4.5 x 25 AA TANDDEM REPEATS.  
 FT DOMAIN 279 347 TWO DIRECTLY REPEATED 27 AMINO ACID  
 FT BLOCKS SEPARATED BY 15 AMINO ACIDS.  
 FT HYDROPHILIC.  
 FT DOMAIN 348 411 GLY/PRO-RICH (CELL WALL-SPANNING).  
 FT DOMAIN 442 448 CONSERVED IN GRAM-POSITIVE COCCI SURFACE  
 FT DOMAIN 449 454 PROTEINS.  
 FT SEQUENCE 483 AA; 53472 MW; 68f87f28db534448 CRC64;  
 Query Match 8.8%; Score 157.5; DB 1; Length 483;  
 Best Local Similarity 24.8%; Pred. No. 0.75;  
 Matches 109; Conservative 68; Mismatches 150; Indels 113; Gaps 24;  
 QY 1 MACNNSG-----KDGASANSADSVKGP-----NLTEISKTTESNAV-----V 40  
 DB 1 MAKNTNRRIVSLRKIKKGTASAVA-LSTIGGLVNTNEVSARFPGTVEENPDKAREL 59  
 QY 41 LAKEVET--LLASIDELATKAIGKIGNGGLEANSQNTSLSGAVASIDLIAEKLAVL 98  
 DB 60 LKRYDVENSMLQANDKLTLE-----NNNL-TDQNNK-----LTTENKMLT 99  
 QY 99 ---KN--EELKIKIDTAQCSTEFNKLKSEHAVIAGLDNDNDAORAILKHKANKDGA 153  
 DB 100 DONKKLTENKMLTQNNKLTLE--NKLKAE-----NRLTTEN--KGLITKLSEAEER 151  
 QY 154 AELEKLFK-AVENLSKAADFTLKNV--KELTS-----PIVHGNSGKD 194  
 DB 152 ANKERENNEALIGTLTKTIDEIVKDKIAEQESKETIGTLTKTIDEIVADKIAEQESKET 211  
 QY 195 GNTSANSADSVKGPNTLEISKTTESNAVLAKEI--ETLLASI----- 238  
 DB 212 IGLTKLTIDEIVK-----DKIAKE--QESKETIGTLTKTIDEIVKDKIAEQESKODIGALK 266  
 QY 239 DELATKAIGKTIQONGGLAVEAGHNGTLLAGAYITISKI--ITOKLDIGLKNSEK----- 289  
 DB 267 QELAKDCGKNVSEASRKGRLRDSDASREAKQVEKDKIANLTAEIDYKKEKQISDSARQ 326  
 QY 290 -LKEKTIENAKKCEDFTKLEGEHAOLGIEVNTDENAKKAILITDAKDGAELEKLFK 348  
 DB 327 GLRDLSDASREAKQVEKALDEANSKILAEKLNKLEESKKLTLEKEK-----AELOAKLE 382  
 QY 349 AVENLAKAKEMLANVYKEL 368  
 DB 383 AE---AKALKLEOLAKQAEEL 399  
 RESULT 13  
 RBP2\_PLAVB  
 ID RBP2\_PLAVB STANDARD; PRT; 1251 AA.  
 AC Q00799;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).  
 GN RBP2.  
 OS Plasmodium vivax (strain Belem).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=31273;  
 RN [1]  
 RP MEDLINE=92315338; PubMed=1617731;  
 RA Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
 RT "A reticulocyte-binding protein complex of Plasmodium vivax

FT mercozoites.";  
 RL Cell 69:1213-1226(1992).  
 CC -I- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
 CC HUMAN RETICULOCYTE CELLS.  
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M88098; AAA29744.1; -  
 KW Malaria; Receptor; Membrane.  
 FT NON\_TER 1 1251  
 FT NON\_TER 1 1251  
 FT SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;  
 Query Match 8.8%; Score 157.5; DB 1; Length 1251;  
 Best Local Similarity 21.3%; Pred. No. 2.1;  
 Matches 102; Conservative 72; Mismatches 147; Indels 157; Gaps 22;  
 QY 9 DGNASANSADSVK-----GPNLTEISKTTESNAVLA-----KEVETLLASIDELAT 58  
 DB 671 EGHGNVAGQVLENIMEIWDENNLSIDLKQATGKNEIEIKIHTSLKNAKAKIILGHVDSA- 729  
 QY 59 KAIGKTI-----GNGLEA--NOSKNTSLSGAVASIDLI 91  
 DB 730 KYVGIKITPELATELLGDAKLKTQAOLKEESKNNVLETEMNSKNTNEIDVHKNIIDAY 789  
 QY 92 AEKLNVLKNEELKEKIDTAQOCST-----EFPNKLKS----- 123  
 DB 790 KVALEILLAH---SPEIDTKQKSSKLEMGNOYLKAVLINQTKNKISSIKSEAVSVK 846  
 QY 124 -----EHAVLG-----IDNLTDDAORAILKHKANKDGAELE 157  
 DB 847 IGVNSKHSLSLITCSKSDSYDNITALEKOTELQNLANSPTQE---KTNNSD---SKLE 900  
 QY 158 KLFKAVENLSKAADFTLKNVKEITSPVHGNSGKGNTSANSADSVKGPNTLEISK 217  
 DB 901 KI-----KTFESLKNLKTLEGE--VNAKASSDNHSHVOSKSPV--NPALSEIEKE 950  
 QY 218 ITESNAVLAKEI-----ETLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAYTI 273  
 DB 951 ETDIDSLNTALDELKKGKRCCEVSRYKLIQDYTKELISDTLE-----INTI 997  
 QY 274 SKLITQKLDGLKNSKIKKEKTIENAKKCEDFTKLEGEHAOLGIE--NVTDENAKKAILI 331  
 DB 998 EKVVAKVLAAYIK--KNEDIVQVLTLEHFNFKOVSNHEPTNFDSKNSSELTKA--V 1053  
 QY 332 TDA-----AKDKG-----AAELEKFKAVENLAKAKER--LANSVK 366  
 DB 1054 TDSKTIITSLKGLVIEEVNTEMTNTLESSAKTEALYLNELKKNKTSLNELTYOTSNEVK 1111  
 RESULT 14  
 MST2\_DROHY  
 ID MST2\_DROHY STANDARD; PRT; 1391 AA.  
 AC Q08656;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE AXONEME-ASSOCIATED PROTEIN MST101(2).  
 GN MST101(2).  
 OS Drosophila hydei (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7224;  
 RN [1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RX MEDLINE=95045538; PubMed=7957199;  
RA Neesen J., Padmanabhan S., Buemmann H.;  
RT "Randomly arranged repeats of a novel highly charged 16-amino-acid  
RT motif representing the major component of the sperm-tail-specific  
RT axonem-associated protein family Dmst101 form extended  
RT alpha-helical rods within the extremely elongated spermatozoa of  
RT Drosophila hydei";  
RL Eur. J. Biochem. 225:1089-1095(1994).  
CC -1- FUNCTION: POSSIBLE STRUCTURAL ROLE IN THE SPERM TAIL.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- TISSUE SPECIFICITY: TESTIS. PRIMARY SPERMATOCYTES AND EARLY  
CC SPERMATIDS.  
CC -1- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.  
CC -1- POLYMORPHISM: LENGTH POLYMORPHISMS EXIST BETWEEN DIFFERENT  
CC STAINS, MOST LIKELY CAUSED BY LENGTH VARIATIONS WITHIN THE TANDEM  
CC REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: X73481; CAAS1876.1; -.  
CC DR PIR: S34154; S34154.  
CC DR HSSP: P01032; IC5A.  
CC DR FLYBase: FBgn0020733; Dmst101(2).  
CC KW Sperm; Repeat; Multigene family; Polymorphism.  
CC FT DOMAIN 332 1268 59 X 16 AA APPROXIMATE TANDEM REPEATS OF  
CC FT SEQUENCE 1391 AA; 159000 MW; 1B2A36BF30F4878 CRC64;  
SQ  
Query Match 8.7%; Score 157; DB 1; Length 1391;  
Best Local Similarity 27.3%; Pred. No. 2.5;  
Matches 109; Conservative 42; Mismatches 190; Indels 58; Gaps 17;  
OY 8 KGNMNSANSADSVGPNLTKTESNAVL-AKREVTLLASIDELATFAICKKIG 66  
DB 285 KKKKKKKK--DEKKKELELEILKQDAEBEAKIRGVYKVK---KKCKKALKKKKKDGG 339  
OY 67 NNGLEANOSKNTSLGAVAIIDLAELKLVNKELEKIDTANOCSTFTNKLKSEHA 126  
DB 340 RKMKEEAEKKKCAALAKKQKEED---EKKCKELAKKKKEADEKKCKE-EAANKKKAAB 395  
OY 127 VLGDNLTDDNORAILKK---HANKDGALEKLFKAVENLSKAAQ-DTLKNVKK--- 179  
DB 336 KKKCKAAKKEKKALEKKCEPAKKEKKAERKKCEELAKNKIKKAAEKKKCKEAAKKK 455  
OY 180 -----ELTSPYVHGNSGKDGNTSANSADSVGPNLTKTESNAVL-NAVVL 226  
DB 456 EAAEKKKCEELAKKIKKAAEKKKCEET-AKKGKVAERKKCEELAKKIKKAAEKKKCKKL 514  
OY 227 AVKEIETLLASIDELATK---AIGKKIQNGGLAVEAGHNGTLLAGAYTTISKLTKORL 281  
DB 515 AKKEKETAEKKCKEAAKRRKKAERKKCEKAAKRRKKAERKKCKEKSAAKREAEKKE 573  
OY 282 DGLKSEKLEKIEKNA--KKCESEDTKKLEGEHAQIGINVTDENAKKKALITD----- 333  
DB 574 --KCKEAKKREKKALEKKCEE--AAKREKVAE--RKCKEELAKKIKKAAEKKKCKE 625  
OY 334 -AAKDKGALEKLFKAVENLSKAAK---EMLANSVKE 367  
DB 626 AAKKKEKAAERKKCEELAKKIKKAAEKKKCKKLAKKREKE 664  
RESULT 15  
MYSN\_DROME STANDARD: PRT: 2017 AA.  
AC 099323;

DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).  
GN ZIP.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo.  
RX MEDLINE=90349606; PubMed=2117279;  
RA Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;  
RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain  
RT transcript: conserved sequences in the myosin tail and differential  
RT splicing in the 5' untranslated sequence";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).  
CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR  
CC CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: M35012; AAA28713.1; -.  
CC DR PIR: A36014; A36014.  
CC DR PIR: B36014; B36014.  
CC DR HSSP: P08799; 1MND.  
CC DR FLYBase: FBgn0005634; zip.  
CC DR InterPro: IPR000048; IQ.  
CC DR InterPro: IPR002928; Myosin\_tail.  
CC DR InterPro: IPR001609; myosin\_head.  
CC DR Pfam: PF00612; IQ.1.  
CC DR Pfam: PF00063; myosin\_head.1.  
CC DR Pfam: PF01576; myosin\_tail.1.  
CC DR PRINTS: PR00193; MYOSINHEAVY.  
CC DR ProDom: PD000355; myosin\_head.1.  
CC DR SMART: SM00015; IQ.1.  
CC DR SMART: SM00242; MYSC.1.  
CC DR PROSITE: PS50096; IQ.1.  
CC KW Myosin; Alternative splicing; Coiled coil; Actin-binding;  
CC ATP-binding; Calmodulin-binding.  
CC FT DOMAIN 1 829  
CC FT DOMAIN 1 859  
CC FT DOMAIN 830 859  
CC FT DOMAIN 886 859  
CC FT NP\_BIND 225 232  
CC FT NP\_BIND 225 232  
CC FT DOMAIN 250 260  
CC FT DOMAIN 682 694  
CC FT DOMAIN 705 727  
CC FT DOMAIN 742 758  
CC FT DOMAIN 1303 2017  
CC FT DOMAIN 1303 1970  
CC FT DOMAIN 1971 2017  
CC FT VARSPLIC MISSING (IN SHORT ISOFORM).  
CC FT SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;  
SQ  
Query Match 8.6%; Score 155.5; DB 1; Length 2017;  
Best Local Similarity 23.6%; Pred. No. 4.3;  
Matches 91; Conservative 65; Mismatches 149; Indels 81; Gaps 16;  
OY 26 LTFEIKSTESNAV-----LAVKEVE---TLLASIDELATFAICKIG----- 66

Db 1108 VADLKQOLNERRVYDEMQAOLAKREBELTQTLIRIDESATKATAQAQBELESQIAEI 1167  
QY 67 NNGLEAQSNTSLISGAYASIDLIAEKLNVLKNELKEKIDTAKOCSTEFNTKXSEHA 126  
1168 QEDLEAKKAAKAK-----AEKVRDLSEELKKNELIDSLDTTAAQ--OELRSKREQELA 1221  
QY 127 VLGIDNLTDDNAQRAIILKKHAN-----KDKGALELEKLFKAVENLSKA-----AOD 172  
Db 1222 TL-----KKSLEETVNHGVLADMHRKHSHOEINSINDOLENLRKAKTYLEKAKG 1271  
QY 173 TLKNAVKELTSPYVHGNNSGKDGNTSANSADSVKGPNTLEISKITESNAVLAKEIE 232  
1272 TLEENADLATELSVNSSSROENDRRKQAESQI-----AEOYKIAEIERARSELQEC 1326  
QY 233 TLASIDELATKAIGKIQONGGLAVEAGHN--GTLLAGAYTISKLTOKIDGLKN----- 286  
Db 1327 TKIQOEAENITNQL-EKAEIKASAAVKSASNMESQITEAQOLEETPROKL-GLSSKLRQ 1384  
287 -----SEKLEKEIENAKKSEDEFTKKEGEHAQGIENVTDENAKKAILITDAKDKGAEE 342  
Db 1385 IESEKALQOLEEDDEKAKRYERKLAEVTTOM-----OEIKKAEEDADLAK-----E 1433  
QY 343 LE---KLFKAVENLAKAKEMLAN 364  
1434 LEEGKKRINKDIEALERQYKEILIAQN 1459

Search completed: March 18, 2002, 10:11:57  
Job time: 975 sec



result No.	Score	Query Match	Length	DB	ID	Description
1	901	50.1	192	2	09RR52	09rr52 borrella bu
2	901	50.1	209	2	044717	044717 borrella bu
3	899	50.0	1	2	09s3p1	09s3p1 borrella bu
4	889	49.4	192	2	031117	031117 borrella bu
5	876	48.7	193	2	p94234	p94234 borrella bu
6	876	48.7	194	2	031122	031122 borrella bu
7	876	48.7	201	2	09R0R8	09rr08 borrella bu
8	863	48.0	191	2	044726	044726 borrella bu
9	841	46.8	185	2	031123	031123 borrella bu
10	826	45.9	178	2	044895	044895 borrella bu
11	823	45.8	178	2	044879	044879 borrella bu
12	739	41.1	178	2	044594	044594 borrella bu
13	653.5	36.3	212	2	044705	044705 borrella bu
14	647.5	36.0	193	2	031115	031115 borrella bu
15	646	35.9	202	2	008831	008831 borrella ta
16	646	35.9	211	2	044720	044720 borrella bu
17	644.5	35.8	192	2	09s3p3	09s3p3 borrella bu
18	644.5	35.8	192	2	09rrb1	09rrb1 borrella bu
19	644.5	35.8	193	2	09RR53	09rr53 borrella bu

20	64.5	35.8	200	2	Q9R0R9	Q9rq9	borrella	bu
21	64.2	35.7	193	2	P94237	P94237	borrella	bu
22	641	35.7	188	2	Q9XDH3	Q9xdh3	borrella	bu
23	640.5	35.6	191	2	Q31120	Q31120	borrella	bu
24	640	35.6	184	2	Q9S504	Q9s504	borrella	bu
25	640	35.6	193	2	P94242	P94242	borrella	bu
26	640	35.6	211	2	Q44977	Q44977	borrella	bu
27	639.5	35.6	192	2	Q9S3P2	Q9s3p2	borrella	bu
28	639.5	35.6	210	2	Q44719	Q44719	borrella	bu
29	637.5	35.5	191	2	P70818	P70818	borrella	bu
30	637.5	35.5	201	2	P96572	P96572	borrella	ja
31	635	35.3	203	2	Q08L38	Q08l38	borrella	bu
32	634.5	35.3	201	2	P96573	P96573	borrella	ja
33	634	35.3	191	2	Q34124	Q34124	borrella	bu
34	630	35.0	194	2	Q31114	Q31114	borrella	bu
35	629.5	35.0	201	2	P96514	P96514	borrella	ja
36	629.5	35.0	201	2	P96516	P96516	borrella	ja
37	629	35.0	200	2	Q9XDH4	Q9xdh4	borrella	bu
38	628.5	35.0	192	2	Q9R7B2	Q9r7b2	borrella	bu
39	627	34.9	211	2	Q44976	Q44976	borrella	bu
40	626	34.8	184	2	Q31121	Q31121	borrella	bu
41	625	34.7	193	2	P94236	P94236	borrella	bu
42	624.5	34.7	203	2	Q50619	Q50619	borrella	bu
43	623.5	34.7	203	2	Q50624	Q50624	borrella	af
44	623.5	34.7	212	2	Q9R1M5	Q9r1m5	borrella	ja
45	618.5	34.4	201	2	P96571	P96571	borrella	ja

## ALIGNMENTS

Query Match	50.1%	Score 901;	DB 2;	Length 192;
Best Local Similarity	98.9%	Pred. No. 2.8e-35;		
Matches 184; Conservative	2;	Mismatches	0;	Indels
			Gaps	0;
QY 1	MACNNSGKDGKNSASNSADESVKGPMLTELSKTTESNAVLAKEVETLLASIDELATPA	60		
DB 7	ISCNNSGKDGKNSASNSADESVKGPMLTELSKTTESNAVLAKEVETLLASIDELATPA	66		
QY 61	IGKTKGNNGLENSQSKTSLSGATAISDLIAEKLNVLNKEELKEKIDPAKOCSTFTTK	120		
DB 67	IGKTKGNNGLENSQSKTSLSGATAISDLIAEKLNVLNKEELKEKIDPAKOCSTFTTK	126		
QY 121	LKSEHVAVLGIDNLVTQDNGRAILKHKANKDGAEELEKLFKAVENTLSKAADTLKNAVKE	180		

Db 127 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 186  
 QY 181 LTSPIV 186  
 Db 187 LTSPIV 192

## RESULT 2

044717  
 ID 044717 PRELIMINARY; PRT; 209 AA.

AC 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
 DE 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 GN OSPC.  
 OC Borrelia burgdorferi (Lyme disease spirochete).  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;

Query Match 50.1%; Score 901; DB 2; Length 209;  
 Best Local Similarity 98.9%; Pred. No. 3.1e-35;  
 Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MACNSGKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 60  
 Db 17 ISCNNSKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 76  
 QY 61 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120  
 Db 77 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 136  
 Db 121 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 180  
 Db 137 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 196  
 QY 181 LTSPIV 186  
 Db 197 LTSPIV 202

## RESULT 3

0953P1  
 ID 0953P1 PRELIMINARY; PRT; 191 AA.

AC 0953P1  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 OC Borrelia burgdorferi (Lyme disease spirochete).  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;

Query Match 49.4%; Score 889; DB 2; Length 192;  
 Best Local Similarity 97.8%; Pred. No. 1e-34;  
 Matches 182; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MACNSGKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 60  
 Db 7 ISCNNSKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 66  
 QY 61 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120  
 Db 67 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 126  
 Db 121 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 180  
 Db 127 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 186

RT Lyme disease Borrelia.";  
 RL Mol. Microbiol. 18:257-269(1995).  
 DR EMBL: I42894; AAB37002.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 191  
 SQ SEQUENCE 191 AA; 20340 MW; 8CA1A64CF17AED8F CRC64;

Query Match 50.0%; Score 899; DB 2; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-35;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CANSKGNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKAIG 62  
 Db 1 CANSKGNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKAIG 60  
 QY 63 KIGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 122  
 Db 61 KIGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120  
 QY 123 SEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 182  
 Db 121 SEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 180  
 QY 183 SPV 186  
 Db 181 SPV 184

## RESULT 4

031117  
 ID 031117 PRELIMINARY; PRT; 192 AA.

AC 031117  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, last sequence update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSPC.  
 OC Borrelia burgdorferi (Lyme disease spirochete).  
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_TaxID=139;  
 RC STRAIN=OC7;  
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF029866; AAB86549.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 192  
 SQ SEQUENCE 192 AA; 20684 MW; 1F0C6BB66E291F6B CRC64;

Query Match 49.4%; Score 889; DB 2; Length 192;  
 Best Local Similarity 97.8%; Pred. No. 1e-34;  
 Matches 182; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MACNSGKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 60  
 Db 7 ISCNNSKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 66  
 QY 61 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120  
 Db 67 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 126  
 Db 121 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 180  
 Db 127 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAADTLKNAVKE 186

OY 181 LTSPiv 186  
 DB 187 LTSPiv 192

RESULT 5  
 P94234 PRELIMINARY: PRT: 193 AA.  
 AC P94234;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RT STRAIN=28354;  
 RT MEDLINE=96296448; PubMed=8709845;  
 RT Lively I., Gibbs C.P., Schuster R.,  
 "Evidence for lateral transfer and recombination in ospC variation in  
 Lyme disease Borrelia".  
 RL MOL. Microbiol. 18:257-269(1995).  
 DR EMBL; L42895; AAB37003.1; -;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 193 AA: 20411 MW: 05B68720F061E2A0 CRC64;

Query Match 48.7%; Score 876; DB 2; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-34;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 248  
 DB 2 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 61  
 OY 249 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308  
 DB 62 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 121  
 OY 309 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368  
 DB 122 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 181

RESULT 6  
 O31122 PRELIMINARY: PRT: 194 AA.  
 AC O31122;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RT STRAIN=OC12;  
 RT Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.,  
 RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF029871; AAB86554.1; -;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.

FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 194 AA: 20640 MW: 00A5E6E2D2CE0F7E CRC64;

Query Match 48.7%; Score 876; DB 2; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-34;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 248  
 DB 10 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 69  
 OY 249 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308  
 DB 70 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 129  
 OY 309 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368  
 DB 130 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 189

RESULT 7  
 O9R0R8 PRELIMINARY: PRT: 201 AA.  
 AC O9R0R8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid cp26.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RT STRAIN=3B6;  
 RX Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;  
 "Population dynamics of a naturally occurring heterogeneous mixture of  
 Borrelia burgdorferi clones".  
 RL Infect. Immun. 67:5709-5716(1999).  
 DR EMBL; AF074465; AAD23912.1; -;  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR Prodom: PD001149; Lipoprotein\_6; 1.  
 KW Plasmid.  
 FT NON\_TER 201  
 FT NON\_TER 201  
 SO SEQUENCE 201 AA: 21457 MW: 806F198295101B07 CRC64;

Query Match 48.7%; Score 876; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-34;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 248  
 DB 20 NSNGKDGNTSANSADSVKGNLTETISKRTTESNAVLAVKEITETLLASIDELATKAIGK 79  
 OY 249 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308  
 DB 80 KIOONGGLAVEAGHNGTLLAGAVTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 139  
 OY 309 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368  
 DB 140 EGEHAOLGIENTVDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 199

RESULT 8  
 O44726 PRELIMINARY: PRT: 191 AA.  
 AC O44726;  
 ID O44726;  
 AC O44726;

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=297;  
RX MEDLINE=94314437; PubMed=8039891;  
RA Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;  
RT "Humoral immune response to outer surface protein C of Borrelia  
burgdorferi in Lyme disease: role of the immunoglobulin M response in  
the serodiagnosis of early infection.";  
RL Infect. Immun. 62:3213-3221(1994).  
DR EMBL: U08284; AAA21460.1; Lipoprotein\_6.  
PF pfam: PF01441; Lipoprotein\_6; 1.  
PR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 191 AA; 20150 MW; C49AA030F0A28717 CRC64;

Query Match 48.0%; Score 863; DB 2; Length 191;  
Best Local Similarity 99.4%; Pred. No. 1.6e-33;  
Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 190 NSGKGNSTANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGK 249  
DB 1 NSGKGNSTANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGK 60  
QY 250 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 309  
DB 61 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 120  
QY 310 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 368  
DB 121 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 179

RESULT 9  
031123 PRELIMINARY; PRT; 185 AA.  
AC 031123;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OC13;  
RX Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF029872; AAB86555.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
PR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 185 AA; 19673 MW; 58DFEE3C7769CAF CRC64;

Query Match 46.8%; Score 841; DB 2; Length 185;  
Best Local Similarity 97.7%; Pred. No. 1.6e-32;  
Matches 172; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 189 NSGKGNSTANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGK 248  
DB 1 NSGKGNSTANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGK 60

DB 10 NSGKGNSTANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGK 69  
QY 249 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 308  
DB 70 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 129  
QY 309 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 364  
DB 130 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 185

RESULT 10  
044995 PRELIMINARY; PRT; 178 AA.  
AC 044995;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MUL;  
RX MEDLINE=95286481; PubMed=7768799;  
RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
RT "Evolution of the Borrelia burgdorferi outer surface protein OspC.";  
RL J. Bacteriol. 177:3036-3044(1995).  
DR EMBL: X84779; CA59250.1;  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
PR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 178 AA; 18911 MW; EEE50CE48EADF1CA CRC64;

Query Match 45.9%; Score 826; DB 2; Length 178;  
Best Local Similarity 98.8%; Pred. No. 7.8e-32;  
Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 196 NTSANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGKIOQNG 255  
DB 1 NTSANSADSVKGNLTETSKKTESNAVLAIVEITLLASIDELATKAIGKIOQNG 60  
QY 256 LAYAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLEGEHAOL 315  
DB 61 LAYAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLEGEHAOL 120  
QY 316 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 368  
DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 173

RESULT 11  
044979 PRELIMINARY; PRT; 178 AA.  
AC 044979;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=272;  
RX MEDLINE=95286481; PubMed=7768799;

RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
RT Hansen K.;  
"Evolution of the Borrelia burgdorferi outer surface protein OspC."  
RL J. Bacteriol. 177:3036-3044(1995).  
DR EMBL: X84785; CAAS9256.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SEQUENCE 178 AA: 18938 MW: 97297473EPE8IE9 CRC64;

Query Match 45.8% Score 823; DB 2; Length 178;  
Best Local Similarity 98.3% Pred. No. 1.1e-31;  
Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 196 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 255  
DB 1 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 60  
256 LAVEAGHNGTLLAGAYTTSKLTOKLDGLKNSKEKIEENAKKCSDEPTKKLEGEHAOL 315  
DB 61 LAVEAGHNGTLLAGAYTTSKLTOKLDGLKNSKEKIEENAKKCSDEPTKKLEGEHAOL 120  
OY 316 GIEVNTDENAKKAILITDAADKGALEKLEKFAVENLAKAKEMLANSVKEL 368  
DB 121 GIEVNTDENAKKAILITDAADKGALEKLEKFAVENLAKAKEMLANSVKEL 173

RESULT 12  
O44994 PRELIMINARY; PRT: 178 AA.

AC O44994;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1P;  
RX MEDLINE=95286481; PubMed=7768799;  
RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
RT Hansen K.;  
"Evolution of the Borrelia burgdorferi outer surface protein OspC."  
J. Bacteriol. 177:3036-3044(1995).  
EMBL: X84782; CAAS9253.1; -  
InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SEQUENCE 178 AA: 18818 MW: 79B435D0837DA36D CRC64;

Query Match 41.1% Score 739; DB 2; Length 178;  
Best Local Similarity 89.0% Pred. No. 8.4e-28;  
Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 196 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 255  
DB 1 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 60  
256 LAVEAGHNGTLLAGAYTTSKLTOKLDGLKNSKEKIEENAKKCSDEPTKKLEGEHAOL 315  
DB 61 LAVEAGHNGTLLAGAYTTSKLTOKIRIENSEKTIKIEENAKKCSDEPTKKLEGEHAOL 120  
OY 316 GIEVNTDENAKKAILITDAADKGALEKLEKFAVENLAKAKEMLANSVKEL 368  
1

DB 121 GIEVNTDENAKKAILITDAADKGALEKLEKFAVENLPOAKEMLANSVKEL 173

RESULT 13  
O44705 PRELIMINARY; PRT: 212 AA.

AC O44705;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN C.  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2591;  
RX MEDLINE=94041630; PubMed=8225587;  
RA Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;  
"Molecular characterization and expression of p23 (OspC) from a North  
American strain of Borrelia burgdorferi."  
Infect. Immun. 61:5097-5105(1993).  
DR EMBL: U01892; AAA16057.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
SEQUENCE 212 AA: 22270 MW: FB2EF2673A384276 CRC64;

Query Match 36.3% Score 653.5; DB 2; Length 212;  
Best Local Similarity 72.4% Pred. No. 9.5e-24;  
Matches 139; Conservative 21; Mismatches 29; Indels 3; Gaps 2;

OY 1 MACNSGKDGKNSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRA 60  
DB 17 ICNNSGKRGKNSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRA 76  
OY 61 ICKKIGNNGLEANSKNSLSGAYAISDLIAEKLNLVKN-BELKKKIDTAKOCSTEFNN 119  
DB 77 IONLIAONLNGANONGSLAGAVIYSLIAEKLGLKNSBELKBIKIDAKCKNAQFND 136  
OY 120 KKKSEHAVIGLDN--LTDNAORAILKHKANKDGALEKLEKFAVENLAKAKEMLANSVKEL 177  
DB 137 KKKSSHAELGIANGAASDANAKAAILKTGTDKGAOLEKLEKFAVENLAKAKEMLANSVKEL 196  
OY 178 VKELTSPYVGN 189  
DB 197 VKELTSPYVAEN 208

RESULT 14  
O31115 PRELIMINARY; PRT: 193 AA.

AC O31115;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
GN OSpC.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OC3;  
RX Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;  
Submitted (OCT-1997) to the EMBL/Genbank/DBD databases.  
DR EMBL: AF029862; AAB6545.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR Prodom: PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:34 ; Search time 118.14 Seconds

(without alignments)  
351.118 Million cell updates/sec

Title: US-09-596-746A-52

Perfect score: 2750  
Sequence: 1 MACNSGKQKQNTSANSANDES.....KAAQALFNSVKELTPVVA 560

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 522463 seqs, 74073290 residues

Number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

\_A\_Geneseq\_1101.\*  
1: /SID58/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID58/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID58/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID58/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID58/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID58/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID58/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID58/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID58/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID58/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID58/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID58/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID58/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID58/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID58/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID58/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID58/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID58/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID58/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID58/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID58/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID58/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2750	100.0	560	22	AAB62724
2	1758	63.9	410	22	AAB62740
3	1754	63.8	386	22	AAB62727
4	1729.5	62.9	384	22	AAB62726
5	1726.5	62.8	408	22	AAB62737
6	1493.5	54.2	373	22	AAB62711
7	1488.5	54.1	397	22	AAB62729
8	1487	54.1	374	22	AAB62710
9	1484	54.0	398	22	AAB62728
10	1474	53.6	400	22	AAB62739
11	1469.5	53.4	401	22	AAB62738

Result No.	Score	Query Match	Length	ID	Description
12	1465	53.3	377	22	AAB62713
13	1460.5	53.1	378	22	AAB62712
14	1452	52.8	378	22	AAB62725
15	1444	52.5	401	22	AAB62733
16	1198.5	43.6	369	22	AAB62714
17	1198.5	43.6	369	22	AAB62716
18	1195.5	43.5	393	22	AAB62731
19	1195.5	43.5	393	22	AAB62732
20	1194	43.4	370	22	AAB62715
21	1191	43.3	394	22	AAB62730
22	1167.5	42.5	368	22	AAB62717
23	1162.5	42.3	367	22	AAB62719
24	1159.5	42.2	391	22	AAB62735
25	1159.5	42.2	391	22	AAB62736
26	1158	42.1	368	22	AAB62718
27	1155	42.0	392	22	AAB62734
28	986.5	35.9	466	16	AAR75740
29	928	33.7	192	22	AAB62721
30	928	33.7	210	16	AAW11935
31	928	33.7	210	16	AAR75727
32	926	33.7	192	15	AAR60886
33	922	33.5	587	16	AAR75746
34	920	33.5	209	22	AAB62720
35	915	33.3	466	16	AAR75739
36	908	33.0	192	22	AAB62703
37	889	32.3	211	22	AAB62722
38	889	32.3	212	16	AAR75729
39	886	32.2	212	18	AAW41824
40	885	32.2	206	22	AAB62723
41	882	32.1	207	16	AAR75730
42	880	32.0	194	15	AAR60896
43	878	31.9	207	18	AAW41823
44	864	31.4	176	15	AAR62772
45	858	31.2	212	12	AAR13140

## ALIGNMENTS

RESULT 1  
AAB62724 standard; Protein: 560 AA.  
XX AAB62724:  
AC  
XX  
XX 03-APR-2001 (first entry)  
DT  
XX  
XX Borrelia sp chimeric ospC protein SEQ ID NO: 52.  
DE  
XX  
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
KW  
XX  
XX Chimeric - Borrelia sp.  
OS  
XX  
XX Chimeric - Borrelia sp.  
PN  
XX  
XX WO200078966-A1.  
PD  
XX  
XX 28-DEC-2000.  
PE  
XX  
XX 19-JUN-2000; 2000MO-US16915.  
PF  
XX  
XX 18-JUN-1999; 99US-0140042.  
PR  
XX  
XX (UYNY) UNIV NEW YORK STATE RES FOUND.  
PA  
XX  
XX (BROO-) BROOK BIOTECHNOLOGIES INC.  
PI  
XX  
XX Datwyler RJ, Seino G, Dykhuzen D, Luft BJ, Gomes-Solecki M;  
DR  
XX  
XX WPI: 2001-050113/06.  
N-PSDB: AAR29028.  
Compositions of ospC polypeptides from strains of Borrelia which cause  
Lyme disease are used to immunize animals and detect immune responses  
to Lyme disease -

XX Claim 43; Page 112-113; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC *Borrelia burgdorferi*, *B. affinis* or *B. garinii*. These can be used as  
CC vaccines against *Borrelia* infection, which is spread by ticks and leads  
CC to Lyme disease.

**SQ** Sequence 560 AA;

Query Match	100.0%	Score 2750	DB 22	Length 560
Best Local Similarity	100.0%	Pred. No. 1.9e-17		
Matches 560	Conservative 0	Mismatches 0	Indels 0	Gaps 0

[illegible]

XX  
PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYN) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPT; 2001-050113/06.

XX

PT Lyme disease are used to immunize animals and detect immune responses

XX  
X

XX

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC *Borrelia burgdorferi*, *B. afzelii* or *B. garinii*. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

**SQ**      **Sequence**      **410 AA;**

Query Match	63.9%	Score 1758;	DB 22;	Length 410;
Best Local Similarity	95.1%	Pred. No. 1.8e-109;		
Matches 366;	Conservative 0;	Mismatches 7;	Indels 12;	Gaps 1;

QY	188	AESPANGSNKGGDSASTPAPESAGPNLTETSKITTSNSMFLVAVKEVELTVSIDE	247
Db	20	aesjcsnsgkqgdasinpadesaqpnltetskitltsnafvavkeveltvside	79
QY	248	LAKKAIQKIDNNNGLUALNNOGSLAGAVAIISTLTTERLSKLNLEELKTEIARAKC	307
Db	80	lakkaigqrtdmnglaalnngsllaagyaistltetksrknoelktetelakakkc	139
QY	308	SEEFTKLKSGBADLKODATDDBAKAAILKTHATTDKGAKFKDFESYBGLLKAQVA	367
Db	140	seeftnklxghadlqgdattddakaaiklthattcgakefkdlfesveglkaaqva	199
QY	368	LITNSVYKRLGH-----RNSGCGDSASTNPDESAGPNLTIVYSKITTSNAPLLA	415
Db	200	ltmsvkeltsjpvvaesppkphmannsgdsasinpdesaqpnltvyskitltsnafiila	259
QY	416	VKEVEALLSIDELSKAIGRRKINDGTLDMEANNESLIGAVEISKLITQKLSVLNSEE	475
Db	260	vkeveallssidelskaigrrkndgtldmeanneesllaagyaieislitqkrlsvlnsee	319
QY	476	LKKRIKEAKDCSOKFTTKLKDSDHAELGIGSVODNNAKKAAILKTHGTROKDGAKELDELFS	535
Db	320	lkkrikeakdcsgkfttkikdshaeljigsvoddnakkaailkthgtkdkgakeldefls	379
QY	536	LESISKAQAQALITNSVKELTNPYVA 560	
Db	380	leisiskaagaaltnsvkeltnpyva 404	







CC to Lyme disease.  
XX  
SQ Sequence 373 AA;

Query Match	54.2%;	Score 1491.5;	DB 22;	Length 373;
Best Local Similarity	82.6%;	Pred. No. 9.3e-92;		
Matches 309;	Conservative 21;	Mismatches 43;	Indels 1;	Gaps 1;

```

Oy      2 ACNNSKQKGNFSANSADSVSGPULTEISKRITDTSNAYLLAVEVEALLSTIDEIAAKI 61
Db      1 acnnsqkgnfnsansadsvsgpnlletsksritdtsnayllaveveallstideiaaki 60
Oy      62 GKRIHONNGLDTEYHNHNGSLAGAYASTLIKOLDLKNEGLKEKEDIIDAAKCKSEFTYTK 121
Db      61 gkrlhngmldteyhnngslagayastllkqlldglkneglekediidaackcsefttk 120
Oy      122 LKEKHTDLGKGCYMDADAKKAILKTNGTKTKGAELCKLFESVEVYSKAAKEMLANSYKE 181
Db      121 lkehtldlgkgymdadackeailktngtktkgaelyklfesvevyskaakemlansyke 180
Oy      182 LTPSPVAESPAMGNSNCKGDSASTNPADSEAKRPNLTEISKRITDTSNAYLLAVEVETL 241
Db      181 ltpspvaespamvngnsnckgdsastnpadseavkgnpletsksritdtsnayllaveketl 239
Oy      242 VLSIDELKRAIKGCRIDNNNLALNNONGSLAGAYASTLIKTEISKRITKLNLEELKTEI 301
Db      240 lvsidelktraklgkridngnglvaengngllagayltsklitqkridglkneklel 299
Oy      302 AKAKKCKSEFTYTKLSGHADCKODATDHAHKAAILKTHAATTDKGAKEFKDFLSEVEGL 361
Db      300 enakkcsefdtkltsghenaglyenvdenakkellltidaackdgaeelekllkvaenla 359
Oy      362 KAAOVALTNSYKEL 375
Db      360 kaakemlansykel 373

```

RESULT	7
AAB62729	

ID AAB62729 standard; Protein; 397 AA.

AC AAB62729;

DT 03-APR-2001 (first entry)

DE Borrella sp chimeric ospc protein SEQ ID NO: 62.

**KW** Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

Chimeric - Borrella sp.  
Chimeric - Borrella sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Sehnost G, Dykhulzen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.  
DR N-PCNP; 2AE30023

DR N-PSDB; AAF29033.

PT Compositions of OspC polypeptides from strains of Borrelia which causes  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -

PS Claim 43; Page 125-126; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.

**SQ** Sequence 397 AA;

Query Match	54.1%;	Score 1488.5;	DB 22;	Length 397;
Best Local Similarity	82.4%;	Pred. No. 1.6e-91;		
Matches 308;	Conservative 22;	Mismatches 43;	Indels 1;	Gaps 1;

[illegible]

RESULT 8

ID AAB62710 standard; Protein; 374 AA

AC AAB62710;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospC protein SEQ ID NO: 24.

Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

05 Chimeric - Borrelia sp.  
05 Chimeric - Borrelia sp.

XX  
XX  
PN

XX  
PD 28-DEC-2000

XX  
PF 19-JUN-2000: 2000WO-US16915

PR 18-JUN-1999; 99US-0140042.

PA (UUNY ) UNIV NEW YORK STATE RES FOUND.

xx  
PI Datwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29014.  
XX Compositions of Ospc polypeptides from strains of Borrelia which cause  
PT Lyme disease are used to immunize animals and detect immune responses  
PT to Lyme disease -  
XX  
XX Claim 43; Page 78-79; 160pp; English.  
PS  
CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX  
XX Sequence 374 AA;  
SO  
Query Match 54.1%; Score 1487; DB 22; Length 374;  
Best Local Similarity 83.4%; Pred. No. 1.9e-91;  
Matches 312; Conservative 20; Mismatches 40; Indels 2; Gaps 2;  
OY 2 ACNNSGKDNSTANSADSVKGNPLTEISKRTDTSNAVLAVKEVALSSIDEIAAKAI 61  
DB 1 acnnsykgdgnsansadesvkgpnltelskkltdsnavlavkevalssideiaakai 60  
OY 62 GKTIHONNGLDTEYNHNGSLAGAYAI STLKOKLDGLKNGELKEKIDAAKCEFTFNK 121  
DB 61 gkthngngldteyngngslagayalstlikgldgkngelkexidaakkcetfnk 120  
OY 122 LKEKHTDGLKRGVTDADAKKAILKTNGTKTGAELKLFESVEVLSKAKEMLANSVKE 181  
DB 121 lkehtldgkrgvtdadakeailkngtkgaeeigklfesvevlskaakemlansvke 180  
OY 182 LTSVVAESPAMGNSGKGSASTNPADESA KGNPLTEISKRTDTSNAVLAVKEVETL 241  
DB 181 ltsvvaespmwmsngkgsdnt -sansadesvkgpnltelskkltesnavlavkevetl 239  
OY 242 VLSIDELAKKAIKGINNNGLAALNNONGSLAGAYAI STLTEKLSKLNLELKEI 301  
DB 240 ltsidela-kalgkklndvsladnengngslagayalstlikklsalidselkaei 298  
OY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDGAKFEKDLFESVEGL 361  
DB 299 ekakckseefaklkghehtldgkegvtddnakkailktndktkgadelleklfesvnl 358  
OY 362 KAAQVALTNSYKEL 375  
DB 359 kaakemltnsvkel 372  
RESULT 9  
AAB62728  
ID AAB62728 standard; Protein: 398 AA.  
XX  
AC AAB62728;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Borrelia sp chimeric ospc protein SEQ ID NO: 60.  
XX  
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
OS Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX  
PN WO200078966-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 19-JUN-2000; 2000MO-US16915.  
XX  
PR 18-JUN-1999; 99US-0140042.  
XX

PA (UYNY ) UNIV NEW YORK STATE RES FOUND.  
PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
XX  
PI Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M,  
XX  
DR WPI; 2001-050113/06.  
DR N-PSDB: AAF29032.  
XX  
XX Claim 43; Page 123; 160pp; English.  
PS  
CC The present invention provides compositions comprising ospc proteins and  
CC chimeric ospc proteins from members of the Borrelia genus. These may be  
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as  
CC vaccines against Borrelia infection, which is spread by ticks and leads  
CC to Lyme disease.  
XX  
XX Sequence 398 AA;  
SO  
Query Match 54.0%; Score 1484; DB 22; Length 398;  
Best Local Similarity 83.2%; Pred. No. 3.2e-91;  
Matches 311; Conservative 21; Mismatches 40; Indels 2; Gaps 2;  
OY 2 ACNNSGKDNSTANSADSVKGNPLTEISKRTDTSNAVLAVKEVALSSIDEIAAKAI 61  
DB 25 acnnsykgdgnsansadesvkgpnltelskkltdsnavlavkevalssideiaakai 84  
OY 62 GKTIHONNGLDTEYNHNGSLAGAYAI STLKOKLDGLKNGELKEKIDAAKCEFTFNK 121  
DB 85 gkthngngldteyngngslagayalstlikgldgkngelkexidaakkcetfnk 144  
OY 122 LKEKHTDGLKRGVTDADAKKAILKTNGTKTGAELKLFESVEVLSKAKEMLANSVKE 181  
DB 145 lkehtldgkrgvtdadakeailkngtkgaeeigklfesvevlskaakemlansvke 204  
OY 182 LTSVVAESPAMGNSGKGSASTNPADESA KGNPLTEISKRTDTSNAVLAVKEVETL 241  
DB 205 ltsvvaespmwmsngkgsdnt -sansadesvkgpnltelskkltesnavlavkevetl 263  
OY 242 VLSIDELAKKAIKGINNNGLAALNNONGSLAGAYAI STLTEKLSKLNLELKEI 301  
DB 264 ltsidela-kalgkklndvsladnengngslagayalstlikklsalidselkaei 322  
OY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDGAKFEKDLFESVEGL 361  
DB 323 ekakckseefaklkghehtldgkegvtddnakkailktndktkgadelleklfesvnl 382  
OY 362 KAAQVALTNSYKEL 375  
DB 383 kaakemltnsvkel 396  
RESULT 10  
AAB62739  
ID AAB62739 standard; Protein: 400 AA.  
XX  
AC AAB62739;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.  
XX  
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
XX  
OS Chimeric - Borrelia sp.  
OS Chimeric - Borrelia sp.  
XX  
PN WO200078966-A1.  
XX

```

PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UWNY ) UNIV NEW YORK STATE RES FOUND.
PA (BR00-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR WPI: 2001-050113/06.
DR N-PDSB: AAF29043.
XX
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 153; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
XX chimeric ospc proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
S0 Sequence 400 AA;

Query Match 53.6%; Score 1474; DB 22; Length 400;
Best Local Similarity 81.4%; Pred. No. 1.5e-90;
Matches 307; Conservative 23; Mismatches 43; Indels 4; Gaps 2;

OY 2 ACNNGSKDGNFSANSGADSVKGPNTLETISKITTDSSNAVLLAKVEEVALSIDETAAKA 61
Db :|||||
25 scnsgkdcgnfsansadesvkgpnltelskkildsnavlakveeallssidelaakai 84
OY 62 GKIKIQQNNGLDTEVYHNHNSLLAGAVAISTLIKQKLIDGLKNBECLKBKIDAAKCSSETFTWK 121
Db 85 gkiklqngngldteyhnngslagayaisrltkqridglneglkexidaakkcssetftnk 144
OY 122 LKEKHDTLQEGCVTDADAKEALIKTNGTKTYKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Db 145 lkehtdlqgevtadadakealiktngctkygaeligkfesvevlksaakemlansvke 204
OY 182 LTSPVVAAS---PAMGSNSGKGGSDASTNPDADESAKGPMLTISKITTDSSNAFVLAVKVEY 238
Db :|||||
205 ltspvaasppkhpsrnmvnnsygdnt--saansadesvkgnltelskkiltesanvvylavkvi 263
OY 239 ETLVTSIDELAKKATIGORTDNNGNLALANNONGSLLAGVAYSTLTITTELSTLKLNLEEK 298
Db 264 etlltsidelakkaigkriqngslaveaghngtllagaystlsklitctqridglnsekik 323
OY 299 TEIAAKKRCSEEFNTKLKSGHADLGODATDDHAANAAILKTHTATTDKGAKEKRDLFEESYE 358
Db 324 etlennakrcsedfktkllegehaqljenvtidenakkailltcaaxdkgaeeleklfkave 383
OY 359 GLLKAQAVALTNYSVKEL 375
Db 384 nlaekaekmlansvkel 400

RESULT 11
AAB62738
ID AAB62738 standard; Protein; 401 AA.
XX
XX AAB62738;
XX
XX AC
XX DT 03-APR-2001 (first entry)
XX
XX Borrella sp chimeric ospc protein SEQ ID NO: 80.
XX
XX Borrella; ospc; Lyme disease; vaccine; chimeric protein; tick.
```

Query Match	53.4%	Score 1469.5	DB 22	Length 401
Best Local Similarity	82.2%	Pred. No. 3e-90		
Matches 310	Conservative 22	Mismatches 40	Indels 5	Gaps 3
Sequence	401 AA			
Query	2	ANNNGSGDONTANSADSEVKGPNLTETISKRTTDSNAVLAVKEVEALSTIDEITAAKAI	61	
Db	25	scmngkgdntcsamsadesvkgpnltetiskrttidsnavllavkeveallssideitaaakai	84	
Qy	62	GKRIHQNNGLDIEYHNHNSLGAVALSTLIKOKLIDGLKNEGLKEKIDAAKCSSETFTNK	121	
Db	85	gkkihngngldceyhngslagayaisltikqldgklneglkheidakkcssetftnk	144	
Qy	122	LKEKHTDLCGCVTDADAKKAILKTNGTKKAEBELGKLFESVEVLSKAKEMLANSVKE	181	
Db	145	lkehndldgkegvtadadakeaalktngtktkgaeelgklfesevlskaakemlansvke	204	
Qy	182	LKSPVAES---PAMGNSNGKGGDSASTPAPDESAAKGPULTETISKRTTDSNAFVLAVKEV	238	
Db	205	lspvaaesppkhpmswnsgkdqnt-sasadesvygpnltetiskrttidsnavvlavke	263	
Qy	239	ETLVLSIDELAKKATGOKIDNNNGLAALNNNGSLGAAVASTITETKLSKLKNEELK	298	
Db	264	etllcsidela-kalgkktkndvslndeedhngslsagyltsnlltkktsaiskdsgekl	322	
Qy	299	TEIAAKKCSSEFTNKLSGADLCQODATDDHAKAAILKTATTDKGAKEFKDLFESVE	358	
Db	323	aetleakkkcsseftaklgehcdlggyvtdnakkailktmndtkgadelaklifesvk	382	
Qy	359	GLIKAAQVALTNSVKEL	375	
Db	383	nlskaakemltnsvkel	399	
RESULT	12			
ID	AAAB62713			
AC	AAAB62713 standard: Protein: 377 AA.			

XX 03-APR-2001 (first entry)  
 DT Borrelia sp chimeric ospc protein SEQ ID NO: 30.  
 XX  
 DE  
 XX  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 (UANY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX  
 DR WPI: 2001-050113/06.  
 DR N-PSDB; AAF29017.  
 XX  
 PT Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 86-87; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 377 AA;

Query Match 53.3%; Score 1465; DB 22; Length 377;  
 Best Local Similarity 81.0%; Pred. No. 5.5e-90;  
 Matches 306; Conservative 23; Mismatches 45; Indels 4; Gaps 2;  
 OY 1 MACNNSGKDGNTSANSADSVKGNLLEISKRTIDSNVLAKEVEALLSSIDEIAAKA 60  
 1 macnnsqkdgntsansadesvkgpnlleinkktdsnavlaveveallssideiaaka 60  
 OY 61 IGKRIHONNGIDTEYNNHNSGLAGAYASTLIKQKLDGKNEGLKEKIDAKKCSFTFN 120  
 61 igkrlhngngldtennhngslagayaistlikqkldgkneglkekidakkcsftfn 120  
 OY 121 KLEKHTDLGEGVTDADAKAEAILKTNGTKGAEEELKLFESVEVLSKAEMLANSVK 180  
 121 kkekhtdlgegvtdadakeailktngtkgaeeelklfesvevlskaemlansvk 180  
 OY 121 kkekhtdlgegvtdadakeailktngtkgaeeelklfesvevlskaemlansvk 180  
 OY 181 ELTSPVAES---PAMGNSGKGDSDASTNPADSAKGNLLEISKRTIDSNVLAKE 237  
 181 eltspvaesppkpsmvnsgkdgt-sansadesvkgpnlleiskrtidsnavlave 239  
 OY 181 eltspvaesppkpsmvnsgkdgt-sansadesvkgpnlleiskrtidsnavlave 239  
 OY 238 VETLVLSIDELAKKAIGOKIDNNNGSLAANNONGSLAGAAISTLTTEKLSKINLEEL 297  
 238 vetlvlsidelakkaigokidnnngslaanngslagaaistltteklskinleel 297  
 OY 238 vetlvlsidelakkaigokidnnngslaanngslagaaistltteklskinleel 297  
 OY 240 IETLISIDELAKKAIGOKIDNNNGSLAANNONGSLAGAAISTLTTEKLSKINLEEL 299  
 240 ietlisisdelakkaigokidnnngslaanngslagaaistltteklskinleel 299  
 OY 298 KTEIAKAKKCSFTFNKLSGHAIDLKQDATDDHAKAAAILKTHATTDKGAKKEFDLFESV 357  
 300 kekienakkcsedfttkllegehaglgientdenakkailltdaakkgaalelktikav 359  
 OY 358 EGLAKAQAVALTNYSKEL 375  
 360 enlakakaemlansvkel 377

RESULT 13  
 AAB62712  
 ID AAB62712 standard; Protein; 378 AA.  
 XX  
 AC AAB62712;  
 XX  
 DT 03-APR-2001 (first entry)  
 XX  
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 28.  
 XX  
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.  
 XX  
 OS Chimeric - Borrelia sp.  
 OS Chimeric - Borrelia sp.  
 XX  
 PN WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-US16915.  
 XX  
 PR 18-JUN-1999; 99US-0140042.  
 XX  
 (UANY ) UNIV NEW YORK STATE RES FOUND.  
 (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX  
 DR WPI: 2001-050113/06.  
 DR N-PSDB; AAF29016.  
 XX  
 PT Compositions of Ospc polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Claim 43; Page 83-84; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospc proteins and  
 CC chimeric ospc proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 378 AA;  
 OY 1 MACNNSGKDGNTSANSADSVKGNLLEISKRTIDSNVLAKEVEALLSSIDEIAAKA 60  
 1 macnnsqkdgntsansadesvkgpnlleinkktdsnavlaveveallssideiaaka 60  
 OY 61 IGKRIHONNGIDTEYNNHNSGLAGAYASTLIKQKLDGKNEGLKEKIDAKKCSFTFN 120  
 61 igkrlhngngldtennhngslagayaistlikqkldgkneglkekidakkcsftfn 120  
 OY 121 KLEKHTDLGEGVTDADAKAEAILKTNGTKGAEEELKLFESVEVLSKAEMLANSVK 180  
 121 kkekhtdlgegvtdadakeailktngtkgaeeelklfesvevlskaemlansvk 180  
 OY 121 kkekhtdlgegvtdadakeailktngtkgaeeelklfesvevlskaemlansvk 180  
 OY 181 ELTSPVAES---PAMGNSGKGDSDASTNPADSAKGNLLEISKRTIDSNVLAKE 237  
 181 eltspvaesppkpsmvnsgkdgt-sansadesvkgpnlleiskrtidsnavlave 239  
 OY 181 eltspvaesppkpsmvnsgkdgt-sansadesvkgpnlleiskrtidsnavlave 239  
 OY 238 VETLVLSIDELAKKAIGOKIDNNNGSLAANNONGSLAGAAISTLTTEKLSKINLEEL 297  
 240 vetlvlsidelakkaigokidnnngslaanngslagaaistltteklskinleel 298  
 OY 298 KTEIAKAKKCSFTFNKLSGHAIDLKQDATDDHAKAAAILKTHATTDKGAKKEFDLFESV 357  
 300 kekienakkcsedfttkllegehaglgientdenakkailltdaakkgaalelktikav 359

Db 299 kaelekakcseeflaklgentldgkvtdnakkalkltnndktkgadelekifeav 358  
 QY 358 EGLKAAQVALTNSVKEL 375  
 Db 359 knlskaakemltinsvkel 376

## RESULT 14

AAB62725 standard; Protein; 378 AA.

AC AAB62725;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 54.

DE Borrelia; ospc: Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYV ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR MPI: 2001-050113/06.

DR N-PSDB; AAF29029.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

PS Claim 43; Page 114-115; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and chimeric ospc proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 378 AA;

Query Match 52.8%; Score 1452; DB 22; Length 378;

Best Local Similarity 58.7%; Pred. No. 4; 1e-89;

Matches 330; Conservative 9; Mismatches 31; Indels 192; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDETIARA 60  
 Db 1 macnsgkdgntsansadesvkgpnlteiskitdsnavllavkeveallssidel-aka 59  
 QY 61 IGGKTHQNGLDTEYNNHNSLAGAVAITLTKOKIDGK-NEGKLEKIDAKKCEETPT 119  
 Db 60 Iggkthngdglneanrnesllagayltstltqkisklsgsegkelaakcseets 119  
 QY 120 NKLKKEHTDLGKEGYTDADAKAELTKTNGT-KTKGAEEIGKLFESYEVLSKAKEWLANS 178  
 Db 120 tkkknhbglytgyvtdenakkaillkanaagkdyveeleklsglskskaakemlans 179  
 QY 179 VKELTSPVVAESPAMGSNSGKGDASSTNPADESAKGNLTETSKITDSNAFLAVKEV 238  
 Db 180 vkeltsppv----- 188

QY 239 ETLVLSTDELAKKAIGOKIDNNNGLAANNQNSLAGAVAITLTETKLSKLNLEELK 298  
 Db 189 ----- 188  
 QY 299 TEIAKAKCSEETPNKLGSHADLGKODATDHAKAAILKTHATDCKAKEFDLFESEV 358  
 Db 189 ----- 188  
 QY 359 GLIKAAQVALTNSVKELGHRNNSGDSASTNPDESAKGNLTETSKITDSNAFLAVKE 418  
 Db 189 ----- 188  
 QY 419 VEALLSSIDELSKAIGKTKITNOGTLNEANRNSLAGAVEISKLITOKLSTVNSSEELK 478  
 Db 231 veallssideiskagiklndgclneanrnesllagayelskltqklsvinseelk 290  
 QY 479 KIEKADCSQKFTTKLSDSHAEIGIOSVODDNKAAILKTHGTRDKGAKELEELFKSLSES 538  
 Db 291 kikeadcsqkfttklkdsnaelgysvqddnakkalkthgtrdkgakeleelfkfsles 350  
 QY 539 LSKAAQALTNSVKELTNPVVA 560  
 Db 351 lskaagaaltinsvkeltnpva 372

## RESULT 15

AAB62733 standard; Protein; 401 AA.

AC AAB62733;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 70.

DE Borrelia; ospc: Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYV ) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR MPI: 2001-050113/06.

DR N-PSDB; AAF29037.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

PS Claim 43; Page 136-137; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and chimeric ospc proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 401 AA;

Query Match 52.5%; Score 1444; DB 22; Length 401;  
 Best Local Similarity 58.5%; Pred. No. 1; 5e-88;

Matches 328; Conservative 10; Mismatches 31; Indels 192; Gaps 4;

Oy	2	ACNNSKRGJNFSANSADSVYGPULTEISKITDTSNAYLLAVKEVALLSIDIEIAKAI	6
Dp	25	scmsgkqgnfsansadsxvqpnllteiskitdsnavyllavkeveallsidel_aka1	83
Oy	62	GKKIHQNNGLDTEYHNHNGSLLAGAYAI STLTKOKLDIK_-NEGLEKEIDAKKCSFTFN	120
Dp	84	gkklindgslidneannmesllagaayltstllitqklsklingseqlkeiaaakcseest	143
Oy	121	KLKEKHNDLGKEBGYTDADAKAIIKTNGT_-KTKAEELGKLFESVEYLSAAKEMLANSV	179
Dp	144	kikdhdhaqlqigvytcdenakkailkanaagdkvgveelklsqslsksaakemlansv	203
Oy	180	KELTSPVYAEESPAGMSGNGKCGDSASTPNDESAKGPNLEISKITDTSNAPVLAYKEVE	239
Dp	204	keltspvy-----	211
Oy	240	TLVLSIDELAKKAIGOKIDNNNGILAANNONGSLLAGAYAI STLITEKLSKLNLEELK	299
Dp	212	-----	211
Oy	300	EIAKAKKCSSEFTNKLKSGHADLKODATDDHAKAAILKTHATTDKAKKEFKDLFESVEG	359
Dp	212	-----	211
Oy	360	LLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTIVTSKITTDSNAFLAYKEY	419
Dp	212	-----hgnsgdsastnpdesakgpnltivtskittdsnaflaykev	254
Oy	420	EALLSIDELSKAIGKKIKNGTLDNEANNRESILAGAYEISKLITOKLVLNSEELKXX	479
Dp	255	eallssidelskaigkikngdtdlneannmesllagayelskltqkvlvnseelkxx	314
Oy	480	IKKAKDSQKFTTKUKDSHAELGIQSVODNNAKKAIIKTHGTDKGAKELEBLFKSLESL	539
Dp	315	ikeakdsqgkfttklkdsaelqigsvodnnaakkaikthgtkdgakeleefkfslesl	374
Oy	540	SKAAQAAITNSVKELTNPVYA	560
Dp	375	skaaqaaaltnsvkeltnpvy	395

Search completed: March 18, 2002, 09:54:35  
Job time: 333 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:46 ; Search time 621.2 Seconds

(without alignments)  
250.303 Million cell updates/sec

Title: US-09-596-746A-52

Perfect score: 2750

Sequence: 1 MACNNSCKDNTSANSADSE.....KAAQALTNVYKELTNPVVA 560

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 27657034 residues

al number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA.Main:\*

1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2750	100.0	560	19	US-09-596-746-52
2	2750	100.0	560	19	US-09-596-746A-52
3	1758	63.9	410	19	US-09-596-746-84
4	1758	63.9	411	19	US-09-596-746A-84
5	1754	63.8	386	19	US-09-596-746-58
6	1754	63.8	386	19	US-09-596-746A-58
7	1734.5	63.1	385	19	US-09-596-746A-56
8	1729.5	62.9	384	19	US-09-596-746-56
9	1726.5	62.8	408	19	US-09-596-746-78

10	1726.5	62.8	409	19	US-09-596-746A-78	Sequence 78, Appl
11	1496.5	54.4	375	19	US-09-596-746A-26	Sequence 26, Appl
12	1492	54.3	374	19	US-09-596-746A-24	Sequence 24, Appl
13	1491.5	54.2	373	19	US-09-596-746-26	Sequence 26, Appl
14	1488.5	54.1	397	19	US-09-596-746-62	Sequence 62, Appl
15	1488.5	54.1	398	19	US-09-596-746A-62	Sequence 62, Appl
16	1487	54.1	374	19	US-09-596-746-24	Sequence 24, Appl
17	1484	54.0	398	19	US-09-596-746-60	Sequence 60, Appl
18	1484	54.0	399	19	US-09-596-746A-60	Sequence 60, Appl
19	1474	53.6	400	19	US-09-596-746-82	Sequence 82, Appl
20	1474	53.6	401	19	US-09-596-746A-82	Sequence 82, Appl
21	1469.5	53.4	401	19	US-09-596-746-80	Sequence 80, Appl
22	1469.5	53.4	402	19	US-09-596-746A-80	Sequence 80, Appl
23	1465	53.3	377	19	US-09-596-746-30	Sequence 30, Appl
24	1465	53.3	377	19	US-09-596-746A-30	Sequence 30, Appl
25	1460.5	53.1	378	19	US-09-596-746-28	Sequence 28, Appl
26	1460.5	53.1	378	19	US-09-596-746A-28	Sequence 28, Appl
27	1452	52.8	378	19	US-09-596-746-54	Sequence 54, Appl
28	1452	52.8	378	19	US-09-596-746A-54	Sequence 54, Appl
29	1444	52.5	401	19	US-09-596-746-70	Sequence 70, Appl
30	1444	52.5	402	19	US-09-596-746A-70	Sequence 70, Appl
31	1203.5	43.8	370	19	US-09-596-746A-32	Sequence 32, Appl
32	1203.5	43.8	370	19	US-09-596-746A-36	Sequence 36, Appl
33	1199	43.6	371	19	US-09-596-746A-34	Sequence 34, Appl
34	1198.5	43.6	369	19	US-09-596-746-32	Sequence 32, Appl
35	1198.5	43.6	369	19	US-09-596-746-36	Sequence 36, Appl
36	1195.5	43.5	393	19	US-09-596-746-66	Sequence 66, Appl
37	1195.5	43.5	393	19	US-09-596-746-68	Sequence 68, Appl
38	1195.5	43.5	394	19	US-09-596-746A-66	Sequence 66, Appl
39	1195.5	43.5	394	19	US-09-596-746A-68	Sequence 68, Appl
40	1194	43.4	370	19	US-09-596-746-34	Sequence 34, Appl
41	1191	43.3	394	19	US-09-596-746A-64	Sequence 64, Appl
42	1191	43.3	395	19	US-09-596-746A-64	Sequence 64, Appl
43	1167.5	42.5	368	19	US-09-596-746-38	Sequence 38, Appl
44	1167.5	42.5	368	19	US-09-596-746A-38	Sequence 38, Appl
45	1167.5	42.5	368	19	US-09-596-746A-42	Sequence 42, Appl

#### ALIGNMENTS

RESULT 1

US-09-596-746-52

Sequence 52, Application US/09596746

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Seino, Gerald

APPLICANT: Dykhuizen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

TITLE OF INVENTION: Borrelia afzelii that Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT APPLICATION NUMBER: US/09/596,746

CURRENT FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 52

LENGTH: 560

TYPE: PRT

ORGANISM: ospc Chimera

US-09-596-746-52

Query Match 100.0%; Score 2750; DB 19; Length 560;

Best Local Similarity 100.0%; Pred. No. 1.1e-180;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MACNNSCKDNTSANSADSEYKGNLTETSKITDSNAVLAVKEVALLSIDEIAKA 60

DB 1 MACNNSCKDNTSANSADSEYKGNLTETSKITDSNAVLAVKEVALLSIDEIAKA 60

```
OY 61 IGGKHONNGDTEYNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
D6 61 IGGKHONNGDTEYNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
OY 121 KLEKHTDGLKEGVTDADAKAEALIKTNGTKTGAEELGKLFESVEVLSSAKAKEMLANSVK 180
D6 121 KLEKHTDGLKEGVTDADAKAEALIKTNGTKTGAEELGKLFESVEVLSSAKAKEMLANSVK 180
OY 181 ELTSPVAAESPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 240
D6 181 ELTSPVAAESPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 240
OY 241 LVLSDIDELAKKAIQOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKTE 300
D6 241 LVLSDIDELAKKAIQOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKTE 300
OY 301 IAKAKKCEFTFNKLSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
D6 301 IAKAKKCEFTFNKLSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
OY 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPMLTVISKITTDSNAFVLAKEVE 420
D6 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPMLTVISKITTDSNAFVLAKEVE 420
OY 421 ALLSSIDELSKAIGKIKINDGTLIDNEANRNESLAGAYEISKLITOKLSVNSEELKKI 480
D6 421 ALLSSIDELSKAIGKIKINDGTLIDNEANRNESLAGAYEISKLITOKLSVNSEELKKI 480
OY 481 KEAKCSOKFTTKLKDISHAELIGISVODDPAKAKAILKTHGTGKGAKELEELFKLSLS 540
D6 481 KEAKCSOKFTTKLKDISHAELIGISVODDPAKAKAILKTHGTGKGAKELEELFKLSLS 540
OY 541 KAAQAAALTNVSKELTNPVYA 560
D6 541 KAAQAAALTNVSKELTNPVYA 560

RESULT 2
US-09-596-746a-52
; Sequence 52, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selhost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 560
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746a-52

Query Match 100.0%; Score 2750; DB 19; Length 560;
Best Local Similarity 100.0%; Pred. No. 1,1e-180;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
D6 61 IGGKHONNGDTEYNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
OY 121 KLEKHTDGLKEGVTDADAKAEALIKTNGTKTGAEELGKLFESVEVLSSAKAKEMLANSVK 180
D6 121 KLEKHTDGLKEGVTDADAKAEALIKTNGTKTGAEELGKLFESVEVLSSAKAKEMLANSVK 180
OY 181 ELTSPVAAESPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 240
D6 181 ELTSPVAAESPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 240
OY 241 LVLSDIDELAKKAIQOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKTE 300
D6 241 LVLSDIDELAKKAIQOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKTE 300
OY 301 IAKAKKCEFTFNKLSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
D6 301 IAKAKKCEFTFNKLSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
OY 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPMLTVISKITTDSNAFVLAKEVE 420
D6 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPMLTVISKITTDSNAFVLAKEVE 420
OY 421 ALLSSIDELSKAIGKIKINDGTLIDNEANRNESLAGAYEISKLITOKLSVNSEELKKI 480
D6 421 ALLSSIDELSKAIGKIKINDGTLIDNEANRNESLAGAYEISKLITOKLSVNSEELKKI 480
OY 481 KEAKCSOKFTTKLKDISHAELIGISVODDPAKAKAILKTHGTGKGAKELEELFKLSLS 540
D6 481 KEAKCSOKFTTKLKDISHAELIGISVODDPAKAKAILKTHGTGKGAKELEELFKLSLS 540
OY 541 KAAQAAALTNVSKELTNPVYA 560
D6 541 KAAQAAALTNVSKELTNPVYA 560

RESULT 3
US-09-596-746-84
; Sequence 84, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selhost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 410
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-84

Query Match 63.9%; Score 1758; DB 19; Length 410;
Best Local Similarity 95.1%; Pred. No. 1,6e-112;
Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;
```

```
OY 1 MACNNSGKDGNTSANSADSVKGNLTETISKITTDSNAFVLAKEVEALLSSIDEIAKA 60
D6 1 MACNNSGKDGNTSANSADSVKGNLTETISKITTDSNAFVLAKEVEALLSSIDEIAKA 60
OY 61 IGGKHONNGDTEYNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
D6 61 IGGKHONNGDTEYNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
```

```
OY 188 AEPSPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 247
D6 20 AEPSPAMGNSGKGDASSTNPADESAAGPMLTEISKITTDSNAFVLAKEVEY 247
OY 248 LAKKAIGOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKEIAKAKC 307
D6 80 LAKKAIGOKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELKEIAKAKC 307
OY 308 SEFTFNKLSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGLKAAQVA 367
```

```

|||||
Db 140 SEFTNKLKSGHADLGDQATDDHAKAAILKTHATTDKGAKKEKDFEVEEGLLKAQVA 199
OY 368 LTNVSKVELCH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 415
Db 200 LTNVSKVELTSPVVAESPCKPRHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 259
OY 416 VKEVEALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSE 475
Db 260 VKEVEALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSE 319
OY 476 LKKIKKADCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 535
Db 320 LKKIKKADCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 379
OY 536 LESLSKAAQALTNVSKVELTNPVVA 560
Db 380 LESLSKAAQALTNVSKVELTNPVVA 404

```

ULT 4

US-596-746A-84  
Sequence 84, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT FILING DATE: 2000-06-19

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 84

LENGTH: 411

TYPE: PRF

ORGANISM: ospC Chimera

US-09-596-746A-84

Query Match 63.9%; Score 1758; DB 19; Length 411;

Best Local Similarity 95.1%; Pred. No. 1.6e-112;

Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

```

188 ASPAMGNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEETLVLSIDE 247
Db 21 ASISGCSNKGSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEETLVLSIDE 80
OY 248 LKKAIKGIKIDNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELTEIAKAKC 307
Db 81 LKKAIKGIKIDNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELTEIAKAKC 140
OY 308 SEFTNKLKSGHADLGDQATDDHAKAAILKTHATTDKGAKKEKDFEVEEGLLKAQVA 367
Db 141 SEFTNKLKSGHADLGDQATDDHAKAAILKTHATTDKGAKKEKDFEVEEGLLKAQVA 200
OY 368 LTNVSKVELCH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 415
Db 201 LTNVSKVELTSPVVAESPCKPRHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 260
OY 416 VKEVEALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSE 475
Db 261 VKEVEALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSE 319
OY 476 LKKIKKADCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 535
Db 321 LKKIKKADCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 380

```

```

OY 536 LESLSKAAQALTNVSKVELTNPVVA 560
Db 381 LESLSKAAQALTNVSKVELTNPVVA 405

```

RESULT 5

US-09-596-746-58

Sequence 58, Application US/09596746

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 58

LENGTH: 386

TYPE: PRF

ORGANISM: ospC Chimera

US-09-596-746-58

Query Match 63.8%; Score 1754; DB 19; Length 386;

Best Local Similarity 95.8%; Pred. No. 2.8e-112;

Matches 364; Conservative 0; Mismatches 4; Indels 12; Gaps 1;

```

OY 193 MGSNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEETLVLSIDE LAKKA 252
Db 1 MGSNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEETLVLSIDE LAKKA 60
OY 253 IGGKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELTEIAKAKCSEFT 312
Db 61 IGGKIDNNNGLAALNNONGSLAGAYASTLITEKLSKLNLEELTEIAKAKCSEFT 120
OY 313 NKLSGSHADLGDQATDDHAKAAILKTHATTDKGAKKEKDFEVEEGLLKAQVALTNSV 372
Db 121 NKLSGSHADLGDQATDDHAKAAILKTHATTDKGAKKEKDFEVEEGLLKAQVALTNSV 180
OY 373 KEIGH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVE 420
Db 181 KEITSPVVAESPCKPRHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVE 240
OY 421 ALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSEELKKI 480
Db 241 ALLSIDELSKAIGKIKINDGTLDNEANRNESLAGAYEISKLTOKLSVANSSEELKKI 300
OY 481 KEAKDCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKSLESLS 540
Db 301 KEAKDCSOKFTTKLDSHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKSLESLS 360
OY 541 KAAQALTNVSKVELTNPVVA 560
Db 361 KAAQALTNVSKVELTNPVVA 380

```

RESULT 6

US-09-596-746A-58

Sequence 58, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

;; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
;; FILE REFERENCE: 2631.1002-001  
;; CURRENT APPLICATION NUMBER: US/09/596,746A  
;; CURRENT FILING DATE: 2000-06-19  
;; PRIOR APPLICATION NUMBER: US 60/140,042  
;; PRIOR FILING DATE: 1999-06-18  
;; NUMBER OF SEQ ID NOS: 84  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 58  
;; LENGTH: 386  
;; TYPE: PRF  
;; ORGANISM: ospc Chimera  
US-09-596-746a-58

Query Match 63.8%; Score 1754; DB 19; Length 386;  
Best Local Similarity 95.8%; Pred. No. 2, 8e-112;  
Matches 364; Conservative 0; Mismatches 4; Indels 12; Gaps 1;

193 MGSNSGSGSASTNPADBSAKGPNITEISKITTSNAFLAVKEVEVLVLSIDELAKKA 252  
1 MGSNSGSGSASTNPADBSAKGPNITEISKITTSNAFLAVKEVEVLVLSIDELAKKA 60  
253 ICGKIDNNGLAALNNONGSLAGAYASTLITEKLSKLNLELTKETIAKAKCSEFT 312  
61 ICGKIDNNGLAALNNONGSLAGAYASTLITEKLSKLNLELTKETIAKAKCSEFT 120  
313 NKLSGSHADIGKODATDHAKAAILKTHATTGKAEKFEFDESVGLKAAQVALTNSV 372  
121 NKLSGSHADIGKODATDHAKAAILKTHATTGKAEKFEFDESVGLKAAQVALTNSV 180  
373 KELGH-----RNSGSGSASTNPDSAKGPNITVTSKITTSNAFLAVKEVE 420  
181 KELTSPVAESPMPKPMANNSGDSASTNPDSAKGPNITVTSKITTSNAFLAVKEVE 240  
421 ALLSIDELSKAIGKIKNDGTLDEANRNEISLAGAYEISLITOKLSVNSEELKKKI 480  
241 ALLSIDELSKAIGKIKNDGTLDEANRNEISLAGAYEISLITOKLSVNSEELKKKI 300  
481 KEAKDSQKFTTKLDSHAELGIQSVODDNAKKAAILKTHGTGKAKAELELFRKSLESLS 540  
301 KEAKDSQKFTTKLDSHAELGIQSVODDNAKKAAILKTHGTGKAKAELELFRKSLESLS 360  
541 KAAQALTNVSKELTNPVVA 560  
361 KAAQALTNVSKELTNPVVA 380

BLT 7  
US-09-596-746a-56  
;; Sequence 56, Application US/09596746A  
;; GENERAL INFORMATION:  
;; APPLICANT: Dattwyler, Raymond J.  
;; APPLICANT: Seinoest, Gerald  
;; APPLICANT: Dykhuisen, Daniel  
;; APPLICANT: Luft, Benjamin J.  
;; APPLICANT: Maria J.C. Gomes-Solecki  
;; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
;; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
;; FILE REFERENCE: 2631.1002-001  
;; CURRENT APPLICATION NUMBER: US/09/596,746A  
;; CURRENT FILING DATE: 2000-06-19  
;; PRIOR APPLICATION NUMBER: US 60/140,042  
;; PRIOR FILING DATE: 1999-06-18  
;; NUMBER OF SEQ ID NOS: 84  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 56  
;; LENGTH: 385  
;; TYPE: PRF  
;; ORGANISM: ospc Chimera  
US-09-596-746a-56

Query Match 63.1%; Score 1734.5; DB 19; Length 385;  
Best Local Similarity 67.0%; Pred. No. 6, 1e-111;  
Matches 375; Conservative 1; Mismatches 3; Indels 181; Gaps 1;

1 MACNNSGKGNISANSADSVKGNITEISKITTSNAFLAVKEVEVLVLSIDELAKKA 60  
1 MACNNSGKGNISANSADSVKGNITEISKITTSNAFLAVKEVEVLVLSIDELAKKA 60  
61 ICKKHONNGLDTEYNNHNSLAGAYASTLITOKLIDGKNEGLKEKIDAAKCSSEFTN 120  
61 ICKKHONNGLDTEYNNHNSLAGAYASTLITOKLIDGKNEGLKEKIDAAKCSSEFTN 120  
121 KKEKHHTDLGKGVYDADAKEAAILKTNCTKTGAEELKLEFESVEVLSKAAEMANSVK 180  
121 KKEKHHTDLGKGVYDADAKEAAILKTNCTKTGAEELKLEFESVEVLSKAAEMANSVK 180  
181 ELTSPVAESPAMGNSGSGSASTNPADBSAKGPNITEISKITTSNAFLAVKEVE 240  
181 ELTSPVAESP----- 191  
241 LVLSIDELAKKAIGKIDNNGLAALNNONGSLAGAYASTLITEKLSKLNLELTKETE 300  
192 ----- 191  
301 IAKAKCSEFTNKLKSGHADIGKODATDHAKAAILKTHATTGKAEKFEFDESVGL 360  
192 ----- 191  
361 LKAAQVALTNVSKELHRNNSGDSASTNPDSAKGPNITVTSKITTSNAFLAVKEVE 420  
192 -----KPFHGNNSGDSASTNPDSAKGPNITVTSKITTSNAFLAVKEVE 239  
421 ALLSIDELSKAIGKIKNDGTLDEANRNEISLAGAYEISLITOKLSVNSEELKKKI 480  
240 ALLSIDELSKAIGKIKNDGTLDEANRNEISLAGAYEISLITOKLSVNSEELKKKI 299  
481 KEAKDSQKFTTKLDSHAELGIQSVODDNAKKAAILKTHGTGKAKAELELFRKSLESLS 540  
300 KEAKDSQKFTTKLDSHAELGIQSVODDNAKKAAILKTHGTGKAKAELELFRKSLESLS 359  
541 KAAQALTNVSKELTNPVVA 560  
360 KAAQALTNVSKELTNPVVA 379

RESULT 8  
US-09-596-746-56  
;; Sequence 56, Application US/09596746  
;; GENERAL INFORMATION:  
;; APPLICANT: Dattwyler, Raymond J.  
;; APPLICANT: Seinoest, Gerald  
;; APPLICANT: Dykhuisen, Daniel  
;; APPLICANT: Luft, Benjamin J.  
;; APPLICANT: Maria J.C. Gomes-Solecki  
;; TITLE OF INVENTION: Groups of Borrelia burgdorferi and  
;; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans  
;; FILE REFERENCE: 2631.1002-001  
;; CURRENT APPLICATION NUMBER: US/09/596,746  
;; CURRENT FILING DATE: 2000-06-16  
;; PRIOR APPLICATION NUMBER: US 60/140,042  
;; PRIOR FILING DATE: 1999-06-18  
;; NUMBER OF SEQ ID NOS: 84  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 56  
;; LENGTH: 384  
;; TYPE: PRF  
;; ORGANISM: ospc Chimera  
US-09-596-746-56

Query Match 62.9%; Score 1729.5; DB 19; Length 384;  
Best Local Similarity 66.9%; Pred. No. 1, 3e-110;  
Matches 374; Conservative 1; Mismatches 3; Indels 181; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 60
QY 62 GKRIHONNGLDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
DB 61 GKRIHONNGLDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 120
QY 122 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 121 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 180
QY 182 LTPSPVAESPAMGSGSGSDASTNPDESAGKPNLTEISKITDSNAVLAVKEVETL 241
DB 181 LTPSPVAESP----- 190
QY 242 VLSIDELAKKAIGOKIDNNGGLAALNNONGSLLAGAYAI STLTEKLSKLNLEELKTEI 301
DB 191 ----- 190
QY 302 AKAKKCEFTFNKLSGSHADLGKODATDHAKAAILKTHATTDKGAKEFPDLPESVEGL 361
DB 191 ----- 190
QY 362 KAAQVALTNSVKELGHRNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 421
DB 191 ----- 190
QY 422 LLSIDELSAIKGKIKINDGTLDEANRNSLJAGAYEISKLITOKLSVNSEELKRIK 481
DB 240 LLSIDELSAIKGKIKINDGTLDEANRNSLJAGAYEISKLITOKLSVNSEELKRIK 299
QY 482 EAKDCSQKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 541
DB 300 EAKDCSQKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 359
QY 542 AAQALTNVSKELTNPYVA 560
DB 360 AAQALTNVSKELTNPYVA 378
```

```
RESULT 9
US-09-596-746-78
; Sequence 78, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuiszen, Daniel
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 408
; TYPE: PRN
; ORGANISM: ospC Chimera
US-09-596-746-78
```

```
Query Match 62.8%; Score 1726.5; DB 19; Length 408;
Best Local Similarity 66.7%; Pred. No. 2,4e-110;
Matches 373; Conservative 2; Mismatches 3; Indels 181; Gaps 1;
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
```

```
DB 25 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 84
QY 62 GKRIHONNGLDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
DB 85 GKRIHONNGLDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 144
QY 122 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 145 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 204
QY 182 LTPSPVAESPAMGSGSGSDASTNPDESAGKPNLTEISKITDSNAVLAVKEVETL 241
DB 205 LTPSPVAESP----- 214
QY 242 VLSIDELAKKAIGOKIDNNGGLAALNNONGSLLAGAYAI STLTEKLSKLNLEELKTEI 301
DB 215 ----- 214
QY 302 AKAKKCEFTFNKLSGSHADLGKODATDHAKAAILKTHATTDKGAKEFPDLPESVEGL 361
DB 215 ----- 214
QY 362 KAAQVALTNSVKELGHRNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 421
DB 215 ----- 214
QY 422 LLSIDELSAIKGKIKINDGTLDEANRNSLJAGAYEISKLITOKLSVNSEELKRIK 481
DB 264 LLSIDELSAIKGKIKINDGTLDEANRNSLJAGAYEISKLITOKLSVNSEELKRIK 323
QY 482 EAKDCSQKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 541
DB 324 EAKDCSQKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 383
QY 542 AAQALTNVSKELTNPYVA 560
DB 384 AAQALTNVSKELTNPYVA 402
```

```
RESULT 10
US-09-596-746a-78
; Sequence 78, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuiszen, Daniel
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 409
; TYPE: PRN
; ORGANISM: ospC Chimera
US-09-596-746a-78
```

```
Query Match 62.8%; Score 1726.5; DB 19; Length 409;
Best Local Similarity 66.7%; Pred. No. 2,4e-110;
Matches 373; Conservative 2; Mismatches 3; Indels 181; Gaps 1;
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 85
QY 62 GKRIHONNGLDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
```

```

Db      86 GKKIHONNGIDTEYNHNSGLAGAVAI STLKQKLDGLNEGKKEKIDAAKCEFTTK 145
Qy      122 LKEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 181
Db      146 LKEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 205
Qy      182 LTPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEETL 241
Db      206 LTPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEETL 215
Qy      242 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTEI 301
Db      216 -----KPEHGNSGSDASTNPDESAGPMLTVISKKITDSNAFVLAKEVEE 215
Qy      302 AKAKKCEFTNKLKSGHADLGKQDATDHA KAILKTHATTDKGAKEFKDLFESVEGL 361
Db      216 -----KPEHGNSGSDASTNPDESAGPMLTVISKKITDSNAFVLAKEVEE 215
Qy      362 KAAQVALTNSVKELGHRNNSGSDASTNPDESAGPMLTVISKKITDSNAFVLAKEVEE 421
Db      216 -----KPEHGNSGSDASTNPDESAGPMLTVISKKITDSNAFVLAKEVEE 264
Qy      422 LLSIDELSKAIGKIKNGTIDNNEANRESLIAGAYEISKITLQKLSVNSEBELKKIK 481
Db      265 LLSIDELSKAIGKIKNGTIDNNEANRESLIAGAYEISKITLQKLSVNSEBELKKIK 324
Qy      482 EAKDSQKFTTKIKDSHAELGIGSVODDANAKKAILKTHGTRKGAKELEFKLSLESLSK 541
Db      325 EAKDSQKFTTKIKDSHAELGIGSVODDANAKKAILKTHGTRKGAKELEFKLSLESLSK 384
Qy      542 AAQVALTNSVKELTNPVVA 560
Db      385 AAQVALTNSVKELTNPVVA 403

RESULT 11
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-26

```

Query Match 54.4%; Score 1496.5; DB 19; Length 374;  
 Best Local Similarity 82.7%; Pred. No. 1.5e-94;  
 Matches 310; Conservative 21; Mismatches 43; Indels 1; Gaps 1;

```

Qy      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAKEVEALLSIDETIAKA 60
Db      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAKEVEALLSIDETIAKA 60
Qy      61 IGKRIHONNGIDTEYNHNSGLAGAVAI STLKQKLDGLNEGKKEKIDAAKCEFTTN 120

```

```

Db      61 IGKRIHONNGIDTEYNHNSGLAGAVAI STLKQKLDGLNEGKKEKIDAAKCEFTTN 120
Qy      121 KLEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 180
Db      121 KLEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 180
Qy      181 ELTSPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEET 240
Db      181 ELTSPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEET 239
Qy      241 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTE 300
Db      240 LLSIDELAKKAIGKIKNGTIDNNEANRESLIAGAYEISKITLQKLSVNSEBELKKIK 299
Qy      301 IAKAKKCEFTNKLKSGHADLGKQDATDHA KAILKTHATTDKGAKEFKDLFESVEGL 360
Db      300 IMAKKCEFTNKLKSGHADLGKQDATDHA KAILKTHATTDKGAKEFKDLFESVEGL 359
Qy      361 IKAQVALTNSVKEL 375
Db      360 AKAKEMLANSVKEL 374

RESULT 12
US-09-596-746a-24
; Sequence 24, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-24

```

Query Match 54.3%; Score 1492; DB 19; Length 375;  
 Best Local Similarity 83.5%; Pred. No. 3e-94;  
 Matches 313; Conservative 20; Mismatches 40; Indels 2; Gaps 2;

```

Qy      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAKEVEALLSIDETIAKA 60
Db      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAKEVEALLSIDETIAKA 60
Qy      61 IGKRIHONNGIDTEYNHNSGLAGAVAI STLKQKLDGLNEGKKEKIDAAKCEFTTN 120
Db      61 IGKRIHONNGIDTEYNHNSGLAGAVAI STLKQKLDGLNEGKKEKIDAAKCEFTTN 120
Qy      121 KLEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 180
Db      121 KLEKHTDLGEGYVDADAKAEAILKTNGTKTGAEEGLKFESEVEVLSKAKEMLANSVK 180
Qy      181 ELTSPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEET 240
Db      181 ELTSPVVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAKEVEET 239
Qy      241 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTE 300
Db      240 LLSIDELAKKAIGKIKNGTIDNNEANRESLIAGAYEISKITLQKLSVNSEBELKKIK 298

```

RESULT 14  
US-09-596-746-62  
; Sequence 62, Application US/09596746

```

RESULT 15
US-09-596-746A-62
: Sequence 62, Application US/09596746A
: GENERAL INFORMATION:
: APPLICANT: Dattwyler, Raymond J.
: APPLICANT: Seinstoel, Gerald
: APPLICANT: Dykhuizen, Daniel
: APPLICANT: Luft, Benjamin J.
: APPLICANT: Maria J.C. Gomes-Solecki
: TITLE OF INVENTION: Groups of Borrelia burgdorferi and
: TITLE OF INVENTION: Borrelia azzellii That Cause Lyme Disease In Humans
: FILE REFERENCE: 2631.1002-001
: CURRENT APPLICATION NUMBER: US/09/596,746A
: PRIOR APPLICATION NUMBER: 2000-06-19
: PRIOR FILING DATE: 1999-06-18
: NUMBER OF SEQ ID NOS: 84
: SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 62  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: ospc Chimera  
US-09-596-746a-62

Query Match 54.1%; Score 148.5; DB 19; Length 398;  
Best Local Similarity 82.4%; Pred. No. 5.8e-94;  
Matches 308; Conservative 22; Mismatches 43; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDNTSANSADSVKGPMLTEISKITDSSNAVLAKEVEALLSSIDEIAAKAI 61
: |||||
Db 26 SCNNSGKDNTSANSADSVKGPMLTEISKITDSSNAVLAKEVEALLSSIDEIAAKAI 85
62 GKTIHONGGLTEYNHNSLAGAIAISTLIKOKIDGLKNEGKLEKIDAAKCCSEFTNK 121
|||
86 GKTIHONGGLTEYNHNSLAGAIAISTLIKOKIDGLKNEGKLEKIDAAKCCSEFTNK 145
122 LKEHTDGLKEGVTDADAKAAILKTNGTKTGAEELGKLFESVEYLSKAKEMLANSYKE 181
|||
Db 146 LKEHTDGLKEGVTDADAKAAILKTNGTKTGAEELGKLFESVEYLSKAKEMLANSYKE 205
182 LTSPVAESPAMGNSGKGDSASTNPADESAAGPMLTEISKITDSSNAVLAKEVEETL 241
|||
Db 206 LTSPVAESPAMVNSGKDNT-SANSADSVKGPMLTEISKITDSSNAVLAKEVEETL 264
242 VLSIDELAKKAIGOKIDNNGLAANNONGSLAGAYAISTLITEKLSKLNLEIKTEI 301
: |||||
Db 265 LASTIDELATKAIGKIOONGSLAEGHNGTLLGAYTISKLTOKLDGLKNSKLEKTI 324
OY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDKAKEFKDLFESVEGL 361
|||
Db 325 ENAKKCEDEFKLEGEHAQGIENVTDENAKKAILLITDAKDKGAELKLFKAVENLA 384
OY 362 KAAQVALTNSYKEL 375
|||
Db 385 KAAKEMLANSYKEL 398
```

Search completed: March 18, 2002, 10:08:47  
Job time: 975 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:35 ; Search time 55.5 Seconds

(without alignments)  
227.060 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750  
Sequence: 1 MACNNSGKRGNTSANSANSADES.....KAAQALITNSVKELTPVVA 560Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	986.5	35.9	466	US-08-235-836C-110	Sequence 110, Appl
2	932	33.9	210	US-08-235-836C-30	Sequence 30, Appl
3	928	33.7	209	US-09-196-293-15	Sequence 15, Appl
4	928	33.7	210	US-08-158-353-3	Sequence 3, Appl
5	922	33.5	588	US-08-235-836C-122	Sequence 122, Appl
6	921	33.5	210	US-08-209-603E-15	Sequence 15, Appl
7	915	33.3	466	US-08-235-836C-107	Sequence 107, Appl
8	889	32.3	212	US-08-235-836C-34	Sequence 34, Appl
9	886	32.2	212	US-08-158-353-4	Sequence 4, Appl
10	882	32.1	207	US-08-235-836C-36	Sequence 36, Appl
11	858	31.2	212	US-09-196-293-11	Sequence 11, Appl
12	858	31.2	212	US-08-209-603E-11	Sequence 11, Appl
13	676	24.6	212	US-08-158-353-2	Sequence 2, Appl
14	661	24.0	212	US-08-031-295-2	Sequence 2, Appl
15	661	24.0	212	US-07-903-580-2	Sequence 2, Appl
16	616.5	22.4	209	US-08-235-836C-32	Sequence 32, Appl
17	188	6.8	2285	US-09-308-375-2	Sequence 2, Appl
18	187	6.8	1164	US-08-923-992A-2	Sequence 2, Appl
19	186	6.8	1786	US-08-923-992A-8	Sequence 8, Appl
20	184	6.7	1164	US-08-923-992A-10	Sequence 10, Appl
21	183.5	6.6	1588	PCT-US93-07261-11	Sequence 11, Appl
22	182.5	6.6	1630	PCT-US93-07261-16	Sequence 16, Appl
23	181.5	6.6	630	US-09-973-462-9	Sequence 9, Appl
24	181	6.6	688	US-09-141-047-8	Sequence 8, Appl
25	179	6.5	1128	US-08-923-992A-6	Sequence 6, Appl
26	177.5	6.5	1098	US-08-923-992A-8	Sequence 8, Appl
27	177.5	6.5	1196	US-08-144-121-4	Sequence 4, Appl

28	177.5	6.5	1196	2	US-08-735-893-4	Sequence 4, Appl
29	177	6.4	1388	2	US-08-685-576-1	Sequence 1, Appl
30	176.5	6.4	3248	1	US-08-353-700-1	Sequence 1, Appl
31	176.5	6.4	3248	5	PCT-US93-16216-1	Sequence 1, Appl
32	175.5	6.4	2482	1	US-08-328-254-6	Sequence 6, Appl
33	174	6.3	740	1	US-08-257-073-5	Sequence 5, Appl
34	173	6.3	1038	4	US-09-541-787-4	Sequence 4, Appl
35	171	6.2	1108	4	US-08-923-992A-4	Sequence 4, Appl
36	171	6.2	1388	2	US-08-685-576-4	Sequence 4, Appl
37	170.5	6.2	1151	3	US-08-840-006-6	Sequence 6, Appl
38	170.5	6.2	1200	3	US-08-840-006-5	Sequence 5, Appl
39	170.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appl
40	170.5	6.2	3111	2	US-08-125-077-4	Sequence 4, Appl
41	167	6.1	1147	3	US-08-470-260-5	Sequence 5, Appl
42	167	6.1	1147	3	US-08-471-491-5	Sequence 5, Appl
43	167	6.1	1147	4	US-08-466-662-5	Sequence 5, Appl
44	167	6.1	2101	1	US-08-195-487-4	Sequence 4, Appl
45	167	6.1	2101	5	PCT-US93-06160-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-235-836C-110  
Sequence 110, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising Botrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 35.9%; Score 986.5; DB 4; Length 466;  
Best Local Similarity 51.1%; Pred. No. 86-63;  
Matches 256; Conservative 42; Mismatches 132; Indels 71; Gaps 17;  
QY 1 MACNNSGKRGNTSANSANSADESVKGPNTLITTSIDNSNAVLAVKEVALLSIDEITAAKA 60

```

Db 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKA 76
OY 61 IGGKTHONNGIDTEYNNHNGSLAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 120
Db 77 IGGKTHONNGIDTEYNNHNGSLAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 136
OY 121 KLEKHTDLDGREGVTDADAKFAILKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 180
Db 137 KLEKHTDLDGREGVTDADAKFAILKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 196
OY 181 ELTSPVVAESPAMGNSGKGDSDASTNPDESAKGPMLTEISKRTIDSNVLLAVKEVET 240
Db 197 ELTSPVVAESPAMGNSGKGDSDASTNPDESAKGPMLTEISKRTIDSNVLLAVKEVET 246
OY 241 LVLSTDLAKKAIGOKIDNNNGSLAALNNNGSLAAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 300
Db 247 LVLSTDLAKKAIGOKIDNNNGSLAALNNNGSLAAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 298
OY 301 LVLSTDLAKKAIGOKIDNNNGSLAALNNNGSLAAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 350
Db 299 GKTLYSKKVTSKDKSSTEKEFNEKGEVSEKIIIRADGTR-----LETTGKSDSGKA 351
OY 351 KDLFEES--VSGLLKAQVALTNSVKELGHRNNSGDSASTNPDESAKGPMLTEISKRTID 408
Db 352 KEVLKGYVLEGITLAKETLV--VKE-----GVVTLKSNISKSGE---VSVELNDTD 398
OY 409 SNAFLAVEVEALLSSIDEIS-KAIGKRIKN-----DGLTNEA--NNESLIGAYEI 460
Db 399 SSA--ATKKTAMNDSTLTISADSKRTKDLVFLDTGTTVOQNTAGTLEGSASEI 455
OY 461 SKLITOKLVLNSELKKNIK 481
Db 456 K-----NLSELKNALK 466

RESULT 2
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248362
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 30:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-30

```

```

Query Match 33.9%; Score 932; DB 4; Length 210;
Best Local Similarity 98.4%; Pred. No. 2,1e-59;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKA 76
OY 61 IGGKTHONNGIDTEYNNHNGSLAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 120
Db 77 IGGKTHONNGIDTEYNNHNGSLAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 136
OY 121 KLEKHTDLDGREGVTDADAKFAILKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 180
Db 137 KLEKHTDLDGREGVTDADAKFAILKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 196
OY 181 ELTSPVVAESP 191
Db 197 ELTSPVVAESP 207

```

```

RESULT 3
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Motz, Manfred
; APPLICANT: Sautscheck, Erwin
; APPLICANT: Wilske, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; FILE OF INVENTION: burgdorferi
; CURRENT APPLICATION NUMBER: US/09/196,293
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-09-196-293-15

```

```

Query Match 33.7%; Score 928; DB 4; Length 209;
Best Local Similarity 98.4%; Pred. No. 4e-59;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVLLAVKEVALLSSIDEIAAKA 76
OY 61 IGGKTHONNGIDTEYNNHNGSLAGAYASTLIKOKLDGLKNEGIKEKIDAAKCCSEFTTN 120

```

Db 77 IGGKIHONNGLDTEENNNGSLACAYAI STLTKKQKLDGKNEGLKEKIDAAKCCSEFTFN 136  
Oy 121 KLEKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 180  
Db 137 KLEKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 196  
Oy 181 ELTSPVAESP 191  
Db 197 ELTSPVAESP 207

RESULT 4  
US-08-158-353-3  
; Sequence 3, Application US/08158353  
; Patent No. 5620862  
; GENERAL INFORMATION:  
; APPLICANT: Padula, Steven J.  
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
; TITLE OF INVENTION: Disease  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Millitia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/158.353  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Alice O.  
; REGISTRATION NUMBER: 33,542  
; REFERENCE/DOCKET NUMBER: UCT93-05  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 210 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; 08-158-353-3

Query Match 33.7%; Score 928; DB 1; Length 210;  
Best Local Similarity 98.4%; Pred. No. 4e-59;  
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MACNNSGKDGNTSANSADSVKGNLITETISKKITDSNAVLAVKEVEALLSSIDEIAAKA 60  
Db 17 ISCNNSGKDGNTSANSADSVKGNLITETISKKITDSNAVLAVKEVEALLSSIDEIAAKA 76  
Oy 61 IGGKIHONNGLDTEYNNHNSLLAGAYAI STLTKKQKLDGKNEGLKEKIDAAKCCSEFTFN 120  
Db 77 IGGKIHONNGLDTEYNNHNSLLAGAYAI STLTKKQKLDGKNEGLKEKIDAAKCCSEFTFN 136  
Oy 121 KLEKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 180  
Db 137 KLEKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 196  
Oy 181 ELTSPVAESP 191  
Db 197 ELTSPVAESP 207

Db 5  
US-08-235-836C-122  
; Sequence 122, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Luft, Benjamin J.  
; TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising  
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235.836C  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; REFERENCE/DOCKET NUMBER: BNL93-28A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729  
; INFORMATION FOR SEQ ID NO: 122:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 588 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-836C-122

Query Match 33.5%; Score 922; DB 4; Length 588;  
Best Local Similarity 93.6%; Pred. No. 4.3e-58;  
Matches 191; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

Oy 4 NNSGKDGNTSANSADSVKGNLITETISKKITDSNAVLAVKEVEALLSSIDEIAAKA 63  
Db 286 NNSGKDGNTSANSADSVKGNLITETISKKITDSNAVLAVKEVEALLSSIDEIAAKA 345  
Oy 64 KIHONNGLDTEYNNHNSLLAGAYAI STLTKKQKLDGKNEGLKEKIDAAKCCSEFTFN 123  
Db 346 KIHONNGLDTEYNNHNSLLAGAYAI STLTKKQKLDGKNEGLKEKIDAAKCCSEFTFN 405  
Oy 124 EKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 183  
Db 406 EKHTDLGREGVTDADAKKAI LKTNGTKTGAEELGKLFESVYLSAAKEMLANSVK 465  
Oy 184 SPVVAESP 207  
Db 466 SPVVAESP 207

RESULT 6  
US-08-209-603E-15  
; Sequence 15, Application US/08209603E  
; Patent No. 6248538  
; GENERAL INFORMATION:  
; APPLICANT: FUCHS, RENATE

```

APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3555
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

```

```

Query Match 33.5%; Score 921; DB 4; Length 210;
Best Local Similarity 97.9%; Pred. No. 1.3e-58;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MACNSKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLAVKEVALLSIDETIAAKA 60
DB 17 ISCNNSKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLAVLEVALLSIDETIAAKA 76
QY 61 IGRKHONNGIDTEYNNHNSLLAGAVASTLIKOKLDGKNEGKKEKIDAAKCSFTFTN 120
DB 77 IGRKHONNGIDTEYNNHNSLLAGAVASTLIKOKLDGKNEGKKEKIDAAKCSFTFTN 136
QY 121 KLEKHTDLGREGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSYK 180
DB 137 KLEKHTDLGREGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSYK 196
QY 181 ELNTPVAESP 191

```

```

DB 197 ELNTPVAESP 207
RESULT 7
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107

```

```

Query Match 33.3%; Score 915; DB 4; Length 466;
Best Local Similarity 98.9%; Pred. No. 1e-57;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 NNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLAVKEVALLSIDETIAAKAIGK 63
DB 276 NNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLAVKEVALLSIDETIAAKAIGK 335
QY 64 KIHONNGIDTEYNNHNSLLAGAVASTLIKOKLDGKNEGKKEKIDAAKCSFTFTNLT 123
DB 336 KIHONNGIDTEYNNHNSLLAGAVASTLIKOKLDGKNEGKKEKIDAAKCSFTFTNLT 395
QY 124 EKHTDLGREGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSYKELT 183
DB 396 AKHTDLGREGVTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKAKEMLANSYKELT 455
QY 184 SPVVAESP 191
DB 456 SPVVAESP 463
RESULT 8
US-08-235-836C-34
; Sequence 34, Application US/08235836C

```

```

ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match 32.2%; Score 886; DB 1; Length 212;
Best Local Similarity 90.7%; Pred. No. 4e-56;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

QY 195 SNSKGGDSASTNPADESAKGPNLTETSKRTTDSNAFVLAKEVEETLVLSIDELAKKAIG 254
:|||||:
Db 20 NNSKGGDSASTNPADESAKGPNLTETSKRTTDSNAFVLAKEVEETLVLSIDELAKKAIG 79

QY 255 QKIDNNNGLALNNONGSLAGAVASLTLETKLSKLNLETKETIAKKAKCSPEFTNK 314
|||||:
Db 80 QKIDNNNGLALNNONGSLAGAVASLTLETKLSKLNLETKETIAKKAKCSPEFTNK 139

QY 315 LKSHADIGKODATDDHAKKAILKTHATTDDGAKKEFKDLFESVEELLKAAQVALTNSYKE 374
|||||:
Db 140 LKSHADIGKODATDDHAKKAILKTHATTDDGAKKEFKDLFESVEELLKAAQVALTNSYKE 199

QY 375 LGHRRNNSGGDSASTNP--DESAKGP 397
|:|:|:|:|
Db 200 L-----TSPVVAESPKRP 212

RESULT 10
US-08-235-836C-36
: Sequence 36, Application US/08235836C
: Patent No. 6248562
: GENERAL INFORMATION:
: APPLICANT: Dunn, John J.
: APPLICANT: Luft, Benjamin J.
: TITLE OF INVENTION: No. 6248562a1 Chimeric Proteins Comprising
: TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
: NUMBER OF SEQUENCES: 144
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brookhaven National Laboratory
: STREET:
: CITY: Upton
: STATE: NY
: COUNTRY: USA
: ZIP: 11973
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-36

Query Match 32.1%; Score 882; DB 4; Length 207;  
Best Local Similarity 99.5%; Pred. No. 7.4e-56;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNPDESAGKGNLTVISKITDTSNAFLAVKEVALLSIDELAKKAIG 438  
DB 20 NNSGDSASTNPDESAGKGNLTVISKITDTSNAFLAVKEVALLSIDELAKKAIG 79  
QY 439 NCGTLDNEANRNESILAGAYEISKLITOKLSVNSEELKKIKREAKDCSOKTTTLKDSH 498  
DB 80 NCGTLDNEANRNESILAGAYEISKLITOKLSVNSEELKKIKREAKDCSOKTTTLKDSH 139  
QY 499 AELGISVODDAAKKAIIKTHGTRKGAKELEELFKSLSELSKAQAALVNSVKELTNPV 558  
DB 140 AELGISVODDAAKKAIIKTHGTRKGAKELEELFKSLSELSKAQAALVNSVKELTNPV 199  
QY 559 VA 560  
DB 200 VA 201

Patent 11  
US-196-293-11  
Sequence 11, Application US/09196293  
Patent No. 6183755  
GENERAL INFORMATION:  
APPLICANT: Fuchs, Renate  
APPLICANT: Motz, Manfred  
APPLICANT: Sautscheck, Erwin  
APPLICANT: Wilske, Bettina  
APPLICANT: Preac-Mursic, Vera  
TITLE OF INVENTION: Active proteins from Borrelia  
FILE REFERENCE: burgdorferi  
CURRENT APPLICATION NUMBER: US/09/196,293  
FILING DATE: 1998-11-19  
EARLIER APPLICATION NUMBER: US 08/209,603  
EARLIER FILING DATE: 1994-03-10  
EARLIER APPLICATION NUMBER: US 07/862,535  
EARLIER FILING DATE: 1992-06-19  
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
EARLIER FILING DATE: 1990-12-21  
EARLIER APPLICATION NUMBER: DE P39 42 728.5  
EARLIER FILING DATE: 1989-12-22  
EARLIER APPLICATION NUMBER: DE P40 18 988.0  
EARLIER FILING DATE: 1990-06-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11  
LENGTH: 212  
TYPE: PRT  
ORGANISM: Borrelia burgdorferi  
US-09-196-293-11

Query Match 31.2%; Score 858; DB 4; Length 212;  
Best Local Similarity 88.8%; Pred. No. 3.9e-54;  
Matches 182; Conservative 2; Mismatches 7; Indels 14; Gaps 2;

QY 195 SNSGDSASTNPDESAGKGNLTVISKITDTSNAFLAVKEVALLSIDELAKKAIG 254  
DB 20 NNSGDSASTNPDESAGKGNLTVISKITDTSNAFLAVKEVALLSIDELAKKAIG 79  
QY 255 OKIDNNGLAALNNONGSLAGAYASTLITEKLSKNEBELTEIAKAKCSEFTNK 314  
DB 80 OKIDNNGLAALNNONGSLAGAYASTLITEKLSKNEBELTEIAKAKCSEFTNK 139  
QY 315 LKSGHADLKODATDDAKKAIIKTHATTDGAKAEFDLFEVGLKKAQVALTNSVKE 374  
DB 140 LKSGHADLKODATDDAKKAIIKTHATTDGAKAEFDLFEVGLKKAQVALTNSVKE 199  
QY 375 LGHRNNSGDSASTNP--DESAGK 397  
DB 200 L-----TSPPVAESPKRP 212

RESULT 12  
US-08-209-603E-11  
Sequence 11, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILSKE, BETTINA  
APPLICANT: PREAC-MURSIC, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SAUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LKR-9217-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212  
TYPE: AMINO ACID

TELEPHONE: 617-861-6240

TELEX: 904136

Db 77 IGNIIAQ-NGINAGANQNGSLAGAYVISTLIAEKLOLKNSEELKEKIEDAKKCNKRAFT 13  
 Oy 120 NKLKEKHTDLC--KEGVTDADAKKEILKTNCTKTRKGAEECLKLESEVUSLKAKEMLAN 177  
 Db 136 DKLKSSHAELCIGANGAADANAKAKAILKTNCTKDKGAOLEKLESEVKNLSKAAQETLNN 195  
 Oy 178 SVKELTSPVNAESP 191  
 Db 196 SVKELTSPVNAENP 209  
 RESULT 14  
 US-08-031-295-2  
 ; Sequence 2, Application US/08031295  
 ; Patent No. 5530103  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LIVERY, Ian  
 ; APPLICANT: DORNER, Freidrich  
 ; TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
 ; TITLE OF INVENTION: PREVENTION OF LYME DISEASE  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington, D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/031,295  
 ; FILING DATE: 19930312  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/903,580  
 ; FILING DATE: 25-JUN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/824,161  
 ; FILING DATE: 22-JAN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/727,245  
 ; FILING DATE: 11-JUL-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 30472/142 IMMU  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELETX: 904136







**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:33 ; Search time 81.02 Seconds

(without alignments)  
441.544 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750  
Sequence: 1 MACNNGKKGKNGTNSANSADES.....KAAQALITNSVKELTPVYA 560

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

num DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	886	32.2	212	US-09-974-992-7
2	878	31.9	207	US-09-974-992-3
3	674.5	24.5	211	US-09-974-992-5
4	234	8.5	691	US-09-748-875-1
5	234	8.5	929	US-09-748-875-60
6	230	8.4	707	US-09-748-875-2
7	228	8.3	589	US-09-748-875-14
8	223.5	8.1	690	US-09-748-875-61
9	223.5	8.1	701	US-09-748-875-62
10	221	8.0	711	US-09-748-875-3
11	218.5	7.9	670	US-09-748-875-63
12	215	7.8	1269	US-09-708-427-19881
13	212	7.7	1881	US-10-032-585-7646
14	212	7.7	1881	US-10-072-851-15590
15	209	7.6	1144	US-09-708-427-15045
16	209	7.6	1304	US-09-708-427-15045
17	209	7.6	1313	US-09-708-427-15044
18	201.5	7.3	1690	US-09-614-150-10224
19	201.5	7.3	1690	US-09-614-150-10311
20	197.5	7.2	805	US-09-708-427-3730
21	197.5	7.2	841	US-09-708-427-3729
22	197.5	7.2	875	US-09-815-242-13080
23	197.5	7.2	875	US-10-072-851-13080
24	197	7.2	6281	US-09-815-242-12996
25	197	7.2	6281	US-10-072-851-12996

26	193.5	7.0	2056	US-09-614-150-4824	Sequence 4824, Ap
27	191	6.9	2344	US-09-815-242-12713	Sequence 12713, A
28	191	6.9	2344	US-10-072-851-12713	Sequence 12713, A
29	189.5	6.9	1014	US-09-708-427-19883	Sequence 19883, A
30	189.5	6.9	1018	US-09-708-427-19882	Sequence 19882, A
31	189	6.9	730	US-09-708-427-3731	Sequence 3731, Ap
32	188.5	6.9	576775	US-08-895-6111-2	Sequence 2, Appl
33	188.5	6.9	576775	US-08-895-6111-2	Sequence 2, Appl
34	188.5	6.9	576775	US-09-895-611D-2	Sequence 2, Appl
35	187	6.8	837	US-09-815-242-5883	Sequence 5883, Ap
36	187	6.8	837	US-10-072-851-5883	Sequence 5883, Ap
37	186	6.8	1489	US-09-614-150-6636	Sequence 6636, Ap
38	186	6.8	1786	US-09-742-096-3	Sequence 3, Appl
39	185.5	6.7	3158	US-09-815-242-12611	Sequence 12611, A
40	185.5	6.7	3158	US-10-072-851-12611	Sequence 12611, A
41	183.5	6.7	2434	US-09-815-242-5835	Sequence 5835, Ap
42	183.5	6.7	2434	US-10-072-851-5835	Sequence 5835, Ap
43	183.5	6.7	5795	US-09-815-242-12610	Sequence 12610, A
44	183.5	6.7	5795	US-10-072-851-12610	Sequence 12610, A
45	182.5	6.6	1725	US-10-037-182-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-974-992-7  
; Sequence 7, Application US/09974992  
; GENERAL INFORMATION:  
; APPLICANT: Mathiesen, Marianne J.  
; APPLICANT: Theisen, Michael  
; APPLICANT: Holm, Arne  
; APPLICANT: Ostergaard, Soren  
; TITLE OF INVENTION: Novel OSCP-derived peptide fragments  
; FILE REFERENCE: 459-666P  
; CURRENT APPLICATION NUMBER: US/09/974, 992  
; PRIOR APPLICATION NUMBER: 09/180, 089  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: PCT/DK97/00203  
; PRIOR FILING DATE: 1997-05-02  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Borrelia afzelii  
US-09-974-992-7

Query Match 32.2%; Score 886; DB 6; Length 212;  
Best Local Similarity 90.7%; Pred. No. 7.9e-49;  
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

QY	195	SNSGKGDSASNPADSKKGRNLTETISKITDSNAFVLAWEVEVLVSTIDELAKKAIG	254
DB	20	NNSGKGDSASNPADSKKGRNLTETISKITDSNAFVLAWEVEVLVSTIDELAKKAIG	79
QY	255	OKIDNNNGLAALNNONGSLAGAVATSTLTETKSLKMLDELKTEIAKAKKCEPFTNK	314
DB	80	OKIDNNNGLAALNNONGSLAGAVATSTLTETKSLKMLDELKTEIAKAKKCEPFTNK	139
QY	315	LKSGHADLKODATDDHAKAAALIKTHATDKGAKKEDLFESVEGLKKAQVALTNSVKE	374
DB	140	LKSGHADLKODATDDHAKAAALIKTHATDKGAKKEDLFESVEGLKKAQVALTNSVKE	199
QY	375	LGRNNSGGDSASTNP--DESAKCP	397
DB	200	L-----TSPVAESPKRP	212

RESULT 2  
US-09-974-992-3

```

: Sequence 3, Application US/09/974992
: GENERAL INFORMATION:
: APPLICANT: Mathiesen, Marianne J.
: APPLICANT: Thelesen, Michael
: APPLICANT: Holm, Arne
: APPLICANT: Ostergaard, Soren
: TITLE OF INVENTION: Novel Oscp-derived peptide fragments
: FILE REFERENCE: 459-666P
: CURRENT APPLICATION NUMBER: US/09/974,992
: CURRENT FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: 09/180,089
: PRIOR FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: PCT/DK97/00203
: PRIOR FILING DATE: 1997-05-02
: NUMBER OF SEQ ID NOS: 40
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 3
: LENGTH: 207
: TYPE: PPT
: ORGANISM: Borrelia garinii
: US-09-974-992-3

```

Query Match 8.5%; Score 234; DB 5; Length 929;  
Best Local Similarity 22.9%; Pred. No. 6.3e-07;  
Matches 144; Conservative 94; Mismatches 260; Indels 132; Gaps 22;

US-09-748-875-60  
Sequence 60, Application US/09748875  
GENERAL INFORMATION:  
APPLICANT: BRILES et al.  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
FILE REFERENCE: 454312-3140  
CURRENT APPLICATION NUMBER: US/09/748,875  
PRIOR FILING DATE: 2000-12-26  
PRIOR APPLICATION NUMBER: 09/298,523  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 60  
LENGTH: 929  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-748-875-60

Query Match 8.5%; Score 234; DB 5; Length 929;  
Best Local Similarity 22.9%; Pred. No. 6.3e-07;  
Matches 144; Conservative 94; Mismatches 260; Indels 132; Gaps 22;

US-09-748-875-60  
Sequence 60, Application US/09748875  
GENERAL INFORMATION:  
APPLICANT: BRILES et al.  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
FILE REFERENCE: 454312-3140  
CURRENT APPLICATION NUMBER: US/09/748,875  
PRIOR FILING DATE: 2000-12-26  
PRIOR APPLICATION NUMBER: 09/298,523  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 60  
LENGTH: 929  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-748-875-60

Query Match 8.4%; Score 230; DB 5; Length 707;  
Best Local Similarity 22.2%; Pred. No. 7.7e-07;  
Matches 139; Conservative 100; Mismatches 264; Indels 122; Gaps 21;

US-09-748-875-2  
Sequence 2, Application US/09748875  
GENERAL INFORMATION:  
APPLICANT: BRILES et al.  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
FILE REFERENCE: 454312-3140  
CURRENT APPLICATION NUMBER: US/09/748,875  
PRIOR FILING DATE: 2000-12-26  
PRIOR APPLICATION NUMBER: 09/298,523  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 707  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-748-875-2

Query Match 8.4%; Score 230; DB 5; Length 707;  
Best Local Similarity 22.2%; Pred. No. 7.7e-07;  
Matches 139; Conservative 100; Mismatches 264; Indels 122; Gaps 21;

US-09-748-875-2  
Sequence 2, Application US/09748875  
GENERAL INFORMATION:  
APPLICANT: BRILES et al.  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
FILE REFERENCE: 454312-3140  
CURRENT APPLICATION NUMBER: US/09/748,875  
PRIOR FILING DATE: 2000-12-26  
PRIOR APPLICATION NUMBER: 09/298,523  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 707  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-748-875-2

Db 553 KLTGKAKLLELSDKIDELDAETAKLEVLQKDAEAGNNVAVYFKEGLEKTTAKKAE 612  
QY 522 KDKGAKLELEFKLSLESLAKAQA 546  
Db 613 LEKAEADLKAVDEPETPAPAPA 637

RESULT 7  
US-09-748-875-14

; Sequence 14, Application US/09748875  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; PRIOR FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-14

Query Match 8.3%; Score 228; DB 5; Length 589;  
Best Local Similarity 23.6%; Pred. No. 8.1e-07;  
Matches 148; Conservative 89; Mismatches 247; Indels 142; Gaps 24;

QY 19 ESWK-GPMLT-----EISKRTDSNAVLAVKEVEALLSIDETAIKAKGKIIF-ONNG 70  
Db 1 EGVRSGNMLVTSSGODISKRYAD-----EVESHLESLIKDVKNKLVQHTQNVG 51  
QY 71 LDTE-----YHNSGLAGAYAIS-----TLIKOKLDLKNEGLEKIDA 110  
Db 52 LITKLEIKKKYLYDLKVNVLSEAELETSKTEKEKLTAFEPFKDTLPTEPEKVAEA 111  
QY 111 AKKSEFTNKLKLEKHTD-----LGKE-GVTDADAKKALIKTNGTKGAELG 158  
Db 112 OKVVEAKKKAEDOKERRRNYPTIYKLTLELTAESDVEVKKAELELVKAKESODEE 171  
QY 159 KLFES-VEVLSK-----AAKEMLANGVELTSPVAES-----PAMGSNG 198  
Db 172 KIKQAEAEVESKQAEATRLKIKITDREBAKRRADAKLEAVEKNVATSEODKPRRRAKRG 231  
QY 199 KGDASATNPADSAKGNLTETSKITDSNAFLAVKEVEETLVLSIDELAKKAIGO-KI 257  
Db 232 VSGELATPPKKENDAKSSD--SSVGEETLPSPSLMANESQTEHRKQVDEYIKKMLSEIQL 290  
QY 258 DNNGGLAALNONGSLAGAYASTILTEKLSKL--NLEEL----- 297  
Db 291 DRRK-----HTQVNNLNINIKLSAIKTKYLYELSVLKENSKEEELSTKKAELTAAEQFKK 345  
QY 298 -----KTEIAKAKKCEEF-----TNKLSGHAIDLKODATDHAKAAL 336  
Db 346 DTLKPEKVVAAEAKKVEAKKKAQKQEDRRNYPTIYKLTLELTAESDVEVKKAELEL 405  
QY 337 LKTHATDCKAKKEFKDLFESVEGLKAAQVALTNSVKEIGHRNNSGDSASTNPDESAGK 396  
Db 406 VKEAENSREBEKIKQAKKEVES--KKAETRLKIKIT--DRKKAEEKARRAESSEK- 459  
QY 397 PNLTVISKRTTDSNAFLAVKEVEALLSIDELSKAIGKIKNDGTLNDANRNESLJG 456  
Db 460 --KAAAKQKVDAEEVAL-----EAKIALELEYVQLEKLEKIDSDSDYIKGELRAP 512  
QY 457 AVYISKLITQKLSVNSSEELKKIKAKDCSQKFTTKLDSHAELIGIOSVODNNAKAIL 516  
Db 513 LQ--SKLDTKRAKLSKLELSDKIDELDAEIAKLEVLQKDA-----EGNNVAVYF 561

QY 517 KTHGTCKGAKLELEFKLSLESLAKA 542  
Db 562 K-EGLEKTTAEKKALEKAEADLKRA 586

RESULT 8  
US-09-748-875-61

; Sequence 61, Application US/09748875  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; PRIOR FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-61

Query Match 8.1%; Score 223.5; DB 5; Length 690;  
Best Local Similarity 22.7%; Pred. No. 1.9e-06;  
Matches 143; Conservative 92; Mismatches 263; Indels 131; Gaps 22;

QY 19 ESWK-GPMLT-----EISKRTDSNAVLAVKEVEALLSIDETAIKAKGKIIF-ONNG 70  
Db 38 EGVRSGNMLVTSSGODISKRYAD-----EVESHLESLIKDVKNKLVQHTQNVG 88  
QY 71 LDTE-----YHNSGLAGAYAIS-----TLIKOKLDLKNEGLEKIDA 110  
Db 89 LITKLEIKKKYLYDLKVNVLSEAELETSKTEKEKLTAFEPFKDTLPTEPEKVAEA 148  
QY 111 AKKSEFTNKLKLEKHTD-----LGKE-GVTDADAKKALIKTNGTKGAELG 158  
Db 149 OKVVEAKKKAEDOKERRRNYPTIYKLTLELTAESDVEVKKAELELVKAKESODEE 208  
QY 159 KLFES-VEVLSK-----AAKEMLANGVELTSPVAES-----PAMGSNG 198  
Db 209 KIKQAEAEVESKQAEATRLKIKITDREBAKRRADAKLEAVEKNVATSEODKPRRRAKRG 268  
QY 199 KGDASATNPADSAKGNLTETSKITDSNAFLAVKEVEETLVLSIDELAKKAIGO-KI 257  
Db 269 VSGELATPPKKENDAKSSD--SSVGEETLPSPSLMANESQTEHRKQVDEYIKKMLSEIQL 327  
QY 258 DNNGGLAALNONGSLAGAYASTILTEKLSKL--NLEEL----- 297  
Db 328 DRRK-----HTQVNNLNINIKLSAIKTKYLYELSVLKENSKEEELSTKKAELTAAEQFKK 382  
QY 298 -----KTEIAKAKKCEEFNKLKSGHAD-----LGKODATDHAKAAL 337  
Db 383 DTLKPEKVVAAEAKKVEAKKKAQKQEDRRNYPTIYKLTLELTAESDVEVKKAELELY 442  
QY 338 KTHATDCKAKKEFKDLFESVEGLKAAQVALTNSVKEIGHRNNSGDSASTNPDESAGK 397  
Db 443 KEAENSREBEKIKQAKKEVES--KKAETRLKIKIT--DRKKAEEKARRAESSEK- 495  
QY 398 NLTIVISKRTTDSNAFLAVKEVEALLSIDELSKAIGKIKNDGTLNDANRNESLJG 457  
Db 496 --KAAAKQKVDAEEVAL-----EAKIALELEYVQLEKLEKIDSDSDYIKGELRAPL 549  
QY 458 YEISKLITQKLSVNSSEELKKIKAKDCSQKFTTKLDSHAELIGIOSVODNNAKAIL 517  
Db 550 Q--SKLDTKRAKLSKLELSDKIDELDAEIAKLEVLQKDAAGNNVAVYFKEGLEKTTAE 607  
QY 518 THTGTDKQAKLELEFKLSLESLAKAQA 546  
Db 608 KKALEKAEADLKAVDEPETPAPAPA 636

RESULT 9  
US-09-748-875-62

; Sequence 62, Application US/09748875  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 62  
; LENGTH: 701  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-62

Query Match 8.1%; Score 223.5; DB 5; Length 701;  
Best Local Similarity 22.7%; Pred. No. 2e-06;  
Matches 143; Conservative 92; Mismatches 263; Indels 131; Gaps 22;

```

OY 19 ESKV-GPMLT-----EISKITDSNAVLAVKEVEALLSIDETIAAKAIGKRIH-QNNG 70
DB 38 EGVRSNNLTVTSSGODISKKYAD-----EVESHLESTLKDVKKLKVQHTQNGV 88
OY 71 LDTE-----YHNHNSLAGAVALS-----TLKQKLDGKNGKLEKIDA 110
DB 89 LITKSEIKKRYVLDKVNVLSEALSTKTEKTEKLPATFQFKDPLPEPEKVAEA 148
OY 111 AKKSEETFNKLEKHTD-----LGKE-GVTDAKAEALIKTNGTKGAEEIG 158
DB 149 OKKVEAKKKADEKQKBRNPTTYKTLELEIAESDVEYKKALELVKVAKASQDEE 208
OY 159 KLFES-VEVLSK-----AAKEMLANVKELTSPVVAES-----PAMGSNSG 198
DB 209 KIKQAEAEVSEKQAEATRLKIKITDREKAKRADAKLEKAEVKNVATSEODKPKRRAKG 268
OY 199 KGGDSASNPDESAGKGNLTFSKKITDSNAFVLAVKEVEALLSIDELAKKAIGQ-KI 257
DB 269 VSGELATPDKKENDAKSSD-SSVGEETLPSPSLMANESQTEHRKDVDEYIKMLSEIQL 327
OY 258 DNNNGIALNONGSLAGAVAI STLITEKLSK---NLEEL----- 297
DB 328 DRRK-----HTQNVNLNITKLSAIKTKYLELVKENSKEELJSTKRAELTAPEQFKK 382
OY 298 ----KTEIAKAKKCSSEFTNKLKSGHAD-----LGKODATDHAKAAIL 337
DB 383 DTLKPEKKYAEAEKKVEEKKKAKKADQKEDRRNPTNTYKTLELEAESDVKKAELELV 442
OY 338 KTHATTDGAKFEKDLFESVEGLLKAQVALTNSVKELGHRNNSGDSASTPDSAKCP 397
DB 443 KEANESNREEKIKQAKKEVES--KKAETRLKTIKT--DRKKAEEKKRRAEESSEK-- 495
OY 398 NLTVSKITDSNAFLAVKEVEALLSIDELSKAIGKRIKIDGTILPANRNESLIAGA 457
DB 466 KAAAKAKQVDAEYAL-----EAKIAELEVEYQRLKEKLEKIDESDSDIYKLEGRAPL 549
OY 458 YEISKLTOKLSVNLSEELKKRIKAKDCSOKFTTKLKDASHAELGIQSVQDDNAKAILK 517
DB 550 Q--SKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAEGNNNVEAFAFKGLEKTTAE 607
OY 518 THGTDKGAKLEELFKLSLELSKAQA 546
DB 608 KKAELKAEADLKRAVDPEETPAPAPA 636
```

RESULT 10

US-09-748-875-3

; Sequence 3, Application US/09748875  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 3  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-3

Query Match 8.0%; Score 221; DB 5; Length 711;  
Best Local Similarity 20.5%; Pred. No. 2.9e-06;  
Matches 131; Conservative 106; Mismatches 258; Indels 144; Gaps 20;

```

OY 12 TSANSADSEVKGPNLTFSKKITDSNAVLAVKEVEALLSIDETIAAKAIG-----KRIH 66
DB 48 TSYNKNANES-----QTEHRK-----AAKQVDEDIKMLSEIOTYIKMLSEIOLDKRNH 96
OY 67 QNN-----GLDTEYNHNSGL--AGAVAI STLIKOKLDGLKNGLEKERTDAKKCS 115
DB 97 TQNVNLMRKLSAIQTKYLELVKLEKSKKEELTSTKTELDAFAEPKKEEELTKKLA 156
OY 116 ET-----FTNKIKEKHTDLGK-----EGVTDAD 138
DB 157 EAKOKAKQKEEDFRNPTNTYKTLELEIAEFDVKVKEAELELVKEEAKPRNEKIKQAK 216
OY 139 AKKALIKTNGT-----KTKGAEEIGLKFESVEGLVSKAAKEMLANSVK----- 180
DB 217 AKVESKKAEPATRLKEIKTEKKAEBEAKKAEBSEKKALEAKOKVDTKQKPKRRAKG 276
OY 181 ---ELTSPVVAESPPAMGSNSGKGGDSASTNPDESAGKGNLTFSKKITDSNAFVLAVKE 237
DB 277 VSGELATPDKKENDAKSSDSVG-----EETLPSPSLMANESQTEHR-----KD 322
OY 238 VETLV---LSIDELAKKAIGQKIDNNNGIAL-----NNONGSLAGAVALS 281
DB 323 VDEYIKKMLSEIOLDRRKHTQNVNLMNLIKLSAIKTKYLELVKENSKEELTSTKRAEL 382
OY 282 TLITEKLSK-----LKNLEELKTEIAKAKKCSSEF-----TNKLSGNHDLGKODA 327
DB 383 TAFPEQFKKOTLKPEKKVAEAEKKVEEAKKAKADQKEDRRNPTNTYKTLELEIAESDV 442
OY 328 TDDHAKAAILKTHATTDGAKFEKDLFESVEGLLKAQVALTNSVKELGHRNNSGDSAS 387
DB 443 KYKKALELVKEEANESNREEKIKQAKKEVES--KKAETRLKTIKT--DRKKAEEAK 497
OY 388 TNPDESAGKPNLTFSKKITDSNAFLAVKEVEALLSIDELSKAIGKRIKIDGTIDNEA 447
DB 498 RRAEESSEK---KAAEAKQVDAEYAL-----EAKIAELEVEYQRLKEKLEKIDESDS 549
OY 448 NNEESLIAGAYELSKLTOKLSVNLSEELKKRIKAKDCSOKFTTKLKDASHAELGIQSVQ 507
DB 550 YKLEGRAPLQ--SKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAEGNNNVEAYF 607
OY 508 DNNAKKAILKTHGTDKGAKLEELFKLSLELSKAQA 546
DB 608 KEGLEKTTAKKAELKAEADLKRAVDPEETPAPAPA 646
```

RESULT 11  
US-09-748-875-63  
; Sequence 63, Application US/09748875  
; GENERAL INFORMATION:

```

; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREOF
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-748-875-63

```

Query Match 7.9%; Score 218.5; DB 5; Length 670;  
 Best Local Similarity 22.8%; Pred. No. 3.8e-06;  
 Matches 142; Conservative 81; Mismatches 248; Indels 151; Gaps 23;

```

QY 12 TSANSADESVKGNLTETSKITDSNAVLAVKEVEALLSIDELIAAKAIG-----KTIH 66
DB 48 TSYNNANES-----OTENRK-----AAQVDEDTKMLSETQYIKRMLSEIQDKRRH 96
QY 67 QNNGLDTEYNHNGSLIAGYAISTLIKQDLKNEGLEKIDA-----AKKSEB-----T 117
DB 97 TON-----VNLNRKL-----SAIQTKLYELRYLKEKSKKEELTSYPTNYTKTLELEIAE 147
QY 118 FNKLKEKHTDGLK-----EGVTDADAKAEALIKTGT-----KTGAEBELGKL 161
DB 148 FVYKKEAELELVKEEAKPRNEKIKOKAKAVESKKAATRLTEIKTERKKAEEAKKKA 207
QY 162 ESVEVLSTAAKEMLANSVK-----ELTSPVAESPAMSGSGGD-----SAS 205
DB 208 ESESEKAAEAKOKYDTKQKPKRRKAGVSGELATPDKKENDAKSSDSVSGEELTSPS 267
QY 206 TNPDESAGKPNLTETSKITDSNAFVLAKEVEETLVISIDELAKKAIGQ-KIDNNGLA 264
DB 268 LMAAESQ-----TEHR-----DVEYIKKMLSEIOLDRK----- 299
QY 265 ALNNGNSLAGAVALSTLITEKLSKLK-----NLEEL-----KT 299
DB 300 -HQNVLNLTKLSAIKTKYELSVLKENSKEELTSKTAELTAAFEQPKDTPLEK 357
QY 300 EIAAKKCESEFTTKLKSAGAD-----LGRQDATDDHAKAAILKTHATTD 344
DB 358 KVAAEAKKVEEAKKAKDQEKEDRRNTPTNTYKTLLEAESDVAKAELELVKEANES 417
QY 345 KGAKEFDLFEVSGLLKAQOVALTNSVKELGHNNNGSGSASTNPDESAGPNLVYISK 404
DB 418 RNEKIKQAKKEVES--KKAETATLEKIKT--DDRKAEEEAARRKAESEK---KAAEAK 470
QY 405 KIIDSNAFLAVKEVEALLSIDELSKAIGKIKINDLTLEANRNRESILAGAVEISKLI 464
DB 471 OKVDAEYAL-----EAKIAELEVEYQRLKEKLEIDESSEDIYIKGLAPAO--SKLD 523
QY 465 TOKLSVNSEIFKKIKKEANDCSOKFTTKLDSHAELGTSVODDNKAKAILKTHGTGDK 524
DB 524 TKRAKLSKLELSKIDELDAETLAEVOLKDAEGNNNNVAYFKEGLEKTTAAKKAELER 583
QY 525 GAKLEBELFLESLSKRAQA 546
DB 584 AEADLKAAVDEPETPAPAPQA 605

```

```

RESULT 12
US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY

```

```

; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION:
; US-09-708-427-19881

```

Query Match 7.8%; Score 215; DB 6; Length 1269;  
 Best Local Similarity 22.1%; Pred. No. 1.5e-05;  
 Matches 155; Conservative 118; Mismatches 255; Indels 172; Gaps 33;

```

QY 13 SANGAD-ESYKGNLTETSKITDSNAVLAVKEVEALLSIDELIAAKAIGKTIHONNGL 71
DB 224 SASADSESKALEFSELSTKES-----AKEMEKMAISLOQ-EIKLEKMSSEK 276
QY 72 DTEYNHNGSLIAGYAISTLIKQDLKNEGLEKIDAACKSEFTNTKKEKHTDGLG- 130
DB 277 EALAKSSAGELAAVQEBELALSKSRLL-LETE---QVSSYEAIDELDTQLEBKRRKASES 331
QY 131 -KEGVY---DADKAEALIKT-----NCTKGAEBELGKLFESVEVLSTAAKEMLANSVK 181
DB 332 FKEELSVLQDLQDTQGLQAKLSEQEGINSKAEEL-KEKELLESLSKQOEELT-RTANE 389
QY 182 LTPVAESPAMSGSGGDSAST-----NPDESAGK-----PILTE 220
DB 390 KLAEVLKEKALANVAEYTSNVAITVEONLEELKESKIDSENFSTDLALISQALSNNSE 449
QY 221 ISKTIID-----SNAYLAVK---EVEYTLVSLIDELAKKAIGKIDNNGLA 269
DB 450 LEKLSLELSEHSEAGSAAATOKNMLELDVYRSSQAEEKSIQIKLETFTFAAEQK 509
QY 270 NCSILAGAVALSTLITEKLSKLKELKTELTAKAKKCESEFTN-----KLSGHA 320
DB 510 NAELEQNLMLQKSSDAREBELKESEKSEIQTALAEVEEKQATTOQMOEYKQKASEL 569
QY 321 DL-----GQDATDDHAKAAILK--THA-----TIDKGAKEFKDL-----FESVEGL 361
DB 570 ELSTIOSARNSELLEEDLRIALQGAHEHEDRANTTHORSIELEGLQSSQSKHEDAEGR 629
QY 362 KAAQVALTNS---VKELGHRNS-----GGDS----- 385
DB 630 KOLELLQTEKRIQLEQVSSLEKKGHETBADSGLYGOVAELQSYLEAFQVSSSLE 689
QY 386 -----ASTNPDESAGPNLVYISKITDS--NAFLAVKEVEALLSI--DELSKAIK-- 435
DB 690 AALNITATENEKELTENLNAVTSKKKLEATVDEYVKSISESNLLESTIRNELNTQGLTE 749
QY 436 KIKND-----GTIDNPA-----NRNSILAGAVEISKLI-----QKLSVNSEELK 477
DB 750 STENDLKAAGLOESVEMERKLSAESELEOKGREIDEATTKMELEALHQSISIDSEHRLQ 809
QY 478 KKIK-----EAKDCQKFTT---KLDSHAELGTSVODDNKAKAILKTHG----- 520
DB 810 KAMEEFTSRDSEASSLFTKLDLBEKIKSYEEOQLAEASGKSSLSKEKLEOTLGLRLAAES 869
QY 521 TKDKGAKLEBEL-IRKLSLSKRAQAAILTNS-----VKEL 554
DB 870 VNEKLEQFDOAKERSLOSSESSEILAEATTNNQLIKIKIOEL 909

```

RESULT 13



US-10-032-585-7646  
; Sequence 7646, Application US/10032585  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032.585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7646  
; LENGTH: 1881  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1881)..(1881)  
; OTHER INFORMATION: X-any amino acid  
; 10-032-585-7646

Query Match 7.7%; Score 212; DB 7; Length 1881;  
Best Local Similarity 21.2%; Pred. No. 4e-05;  
Matches 135; Conservative 115; Mismatches 244; Indels 144; Gaps 23;

5 NSGKDGNTSANSADSEVGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKAGK 64  
DB 897 NTKKYELOINNLS-----NNEFKOKINELS-----KKIESLTED-NKFNAKOLEK 944  
QY 65 IHONNGLOTEYNNH--GSLAGAVAIPTLIKOKLDGKNEGLKEKIDAKKCEFTFNK 121  
DB 945 LR-----DTEENNEHMDLRSASVAYNDLKKAKSES-----EETVAKKELELTFSK 993  
QY 122 -----LKEKHTDGLK-----EGVTDA-----DAKEALIKTNGTKTGAELGLK 160  
DB 994 IDNLEKELKEQSKNELEGOLNTDSTNEKFKELEDELKSIKKNKISSONSEL--- 1050  
QY 161 FESVEVLSAAKEMLAN-----SVKELSPVVAESPANGSGKGDASTPAPDE 211  
DB 1051 ---IOLKTEKEDLOAKDEIDKLAETKSNIDNLNSELSSIOSKLEKEEESHSTKDEH 1107  
QY 212 SAKGNPLTEISKITDSNAFVL-----AVKEVEETLVLSIDELAKKAIGOK--- 256  
DB 1108 SLSLENLKLKEEYENTKTSMAKLSAKIEHKKATDEIETKHTHTDLOEHAOKOSOF 1167  
QY 257 -IDNNGGLAALNONGSLAGAVAIPTLITEKL---SKKNLEELKTEIATAKKCSEFT 312  
DB 1168 ESEERNDIKSNLDEANKELSDNREKLSNLEKTELENNKLKTOEKISDLETSAIASEDS 1227  
QY 313 NKLKSGHADLGRQDADTDOHAKAAILKTHATTDKAKE-----FKDLFESVEGLLKA 364  
DB 1228 KSLKHDIEDLKEKIKLE---TTLKENEETMFEKKEQLOVYNDKCKEELACIKLTETK 1283  
QY 365 QVALTNSVYELGHRNNSGGDSASTNPDESAGKPNLTV-ISKRTITDSNAFL----- 414  
DB 1284 EKEINDLIRKL-----EAAKSDHPTERRKLSLIEDTKSESEKNVYIKLNDQIEKLG 1335  
QY 415 -AVKEVEALLSI-----DELSKAIGKKIKINDGTLONEANRNESLILGAVEISKLTITQ 466  
DB 1336 EREKEVVRIOLOLAKTDMEKIKTTLDVLKES--DLEKTKNKSVDTLKKEV----- 1387  
QY 467 KLSVNLSEELKKIKIEAKDCSQFTTKLDSHAELGIGOSVODDNKAKATLKHGTDKGA 526  
DB 1388 -----ENLKEISLLEQOKDPTTKYKELAOLETKSNIDSTTMELEKTELELKV 1440  
QY 527 KLEBELFKSLESIS-----RAAQALATNSYKEL 554  
DB 1441 NELTEATSELTKLODNNGSLTPEIEKTKAKALTKSSKDL 1478

RESULT 14  
US-10-072-851-15590  
; Sequence 15590, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit  
; FILE REFERENCE: ELITRA.028A  
; CURRENT APPLICATION NUMBER: US/10/072.851  
; CURRENT FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: 60/267,636  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15590  
; LENGTH: 1881  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1881)..(1881)  
; OTHER INFORMATION: X-any amino acid  
US-10-072-851-15590

Query Match 7.7%; Score 212; DB 7; Length 1881;  
Best Local Similarity 21.2%; Pred. No. 4e-05;  
Matches 135; Conservative 115; Mismatches 244; Indels 144; Gaps 23;

5 NSGKDGNTSANSADSEVGPNTLEISKTTDSNAVLAVKEVEALLSIDETAAKAGK 64  
DB 897 NTKKYELOINNLS-----NNEFKOKINELS-----KKIESLTED-NKFNAKOLEK 944  
QY 65 IHONNGLOTEYNNH--GSLAGAVAIPTLIKOKLDGKNEGLKEKIDAKKCEFTFNK 121  
DB 945 LR-----DTEENNEHMDLRSASVAYNDLKKAKSES-----EETVAKKELELTFSK 993  
QY 122 -----LKEKHTDGLK-----EGVTDA-----DAKEALIKTNGTKTGAELGLK 160  
DB 994 IDNLEKELKEQSKNELEGOLNTDSTNEKFKELEDELKSIKKNKISSONSEL--- 1050  
QY 161 FESVEVLSAAKEMLAN-----SVKELSPVVAESPANGSGKGDASTPAPDE 211  
DB 1051 ---IOLKTEKEDLOAKDEIDKLAETKSNIDNLNSELSSIOSKLEKEEESHSTKDEH 1107  
QY 212 SAKGNPLTEISKITDSNAFVL-----AVKEVEETLVLSIDELAKKAIGOK--- 256  
DB 1108 SLSLENLKLKEEYENTKTSMAKLSAKIEHKKATDEIETKHTHTDLOEHAOKOSOF 1167  
QY 257 -IDNNGGLAALNONGSLAGAVAIPTLITEKL---SKKNLEELKTEIATAKKCSEFT 312  
DB 1168 ESEERNDIKSNLDEANKELSDNREKLSNLEKTELENNKLKTOEKISDLETSAIASEDS 1227  
QY 313 NKLKSGHADLGRQDADTDOHAKAAILKTHATTDKAKE-----FKDLFESVEGLLKA 364  
DB 1228 KSLKHDIEDLKEKIKLE---TTLKENEETMFEKKEQLOVYNDKCKEELACIKLTETK 1283  
QY 365 QVALTNSVYELGHRNNSGGDSASTNPDESAGKPNLTV-ISKRTITDSNAFL----- 414  
DB 1284 EKEINDLIRKL-----EAAKSDHPTERRKLSLIEDTKSESEKNVYIKLNDQIEKLG 1335

QY 415 -AVEVEALLSSI-----DELSKAIGKKIKNDGTLIDNEANNEISLAGAIEISKLITO 466  
 Db 1336 EREBEVADIOSQLAKTIDMEKIKITIDKYLKES--DIETKTESVDLTKEV----- 1387  
 QY 467 KLVLSNEELKKIKKAKDCSOKFTTKLNDSHAELGISOVDONAKKAILKTGTDKGA 526  
 Db 1388 -----ENIKKEISLIEDQKDDTKRYKELAQLEFKTSNDSTTMELEKTELEKVR 1440  
 QY 527 KELEELFKSLESIS-----KAAQAALTNYSKEL 554  
 Db 1441 NEUTEATSELTKLQDNNOSLTEIEIKTKALTKRSSKD 1478

RESULT 15  
 US-09-708-427-15046  
 Sequence 15046, Application US/09708427  
 APPLICANT: N. ALEXANDROV et al.  
 TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
 FILE REFERENCE: 2750-1243P  
 CURRENT APPLICATION NUMBER: US/09/708,427  
 CURRENT FILING DATE: 2000-11-09  
 NUMBER OF SEQ ID NOS: 85364  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 15046  
 LENGTH: 1144  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: 1..1144  
 OTHER INFORMATION: Xaa is any amino acid  
 NAME/KEY: misc.feature  
 LOCATION: 1..1144  
 OTHER INFORMATION: Ceres Seq. ID 1828629  
 US-09-708-427-15046

Query Match 7.68; Score 209; DB 6; Length 1144;  
 Best Local Similarity 21.8%; Pred. No. 3.1e-05;

Matches 155; Conservative 116; Mismatches 247; Indels 192; Gaps 31;

QY 2 ACNNSG-----KDNSTANSADSEVKGPNLTIEISKKITDSNAVL----- 40  
 Db 102 SCTNSSVBEMKNKYHELEKEVEESNRKSSASSEME-----SVKKQLAEIHNHYLHETKSD 156

QY 41 -----IAKVEEALLSIDEIA-----AKAIGKKI-----HONNG 70  
 Db 157 NAAOKRIELEKTIETAOFTDLEFGROYCIAKEASKLENTLVESIKSELEISOEKTRA 216

QY 71 LDTEYNNGSLLAGAIAIITLIKOKLD-GLKNGLKEDIKAKKCSFTTNKLKERTDL 129  
 Db 217 LDNE-----KAATSNIONLDORTLSTIELERCKVEEESKMDMLTALQEASTE- 268

QY 130 GKEGVTDADAKKALIKTNGTKTKGAEEIGLFESEVYLKSAKE-----ML--ANSY 179  
 Db 269 -----SSEAKATLLVQ-----EELKNCESQVDSLAKLAKETNEKYEKMLEDARNEI 315

QY 180 KELTSPYVA-ESPAMGNSG-----KGDASATNPADESAKGPNLTIEIS----- 222  
 Db 316 DSLKSTYDSIONFEENSKAMGEQELHMGCVKKESEBENSSQOEVSRLVNLKSESEDA 375

QY 223 -----KKTDSNFAVLKVEFTLVLSIDE-----LAKKAIGOKIDNNN--GLAALN 267  
 Db 376 CAKKEERASLKNMLKVAEGEVKTLQETLGAKAESMKLKEESLDKEDLKNVTAIEISSLR 435

QY 268 NONGSLLAGAIAISTL--ITEKLSIKLNL--EELKTEIAKAKKCSFTTNKLKSGHA 320  
 Db 436 EMEGSYLEKEIEELSKVESLVDEKETLQISITQEAELKGREAHMKOIE-----LSTANA 491

QY 321 DLGKQDATDTHAK--AAILKTHATTDKGA--KEFKDLFESEVGL----- 360

Db 492 SL-----VDEATKIQISIOESEDLEKEKEAGYLKIEELSVANESLADNVTDLOSIOESK 546  
 QY 361 -LKAQVALTNYSKELGHRNNGGDSAS--TNPDSA-----KGPNTLVISK 405  
 Db 547 DLKEEVAYLKIEELSVANESLVDKETKLOHIDQEAELGREGASHLKKIEELSKENEN 606

QY 406 ITDSNAFLAVKEV-----ALLSIDELSKAIGKKIKNDGTLIDNEANRNE----- 451  
 Db 607 IVDNVAANNQNTAEEKDLREBEVAYLKIDELSTANGTLADNVTNLONISEENKELRE 666

QY 452 -SLINGAIEISKLTOKLSVYNSEELKKIKKAKDCSOKFTTKLDSHAELGISOVDON 510  
 Db 667 TLLKKAELSEL--NESLVDKASKLQTVQOENELREBEVAYLKIEELSKLHEITLSDQ 724

QY 511 AKKAILKTHGTGDKAKLEELF-KSLESLSKAAQAALTNYSKELTNPVY 559  
 Db 725 FTKIQISNHEKE--LKERETAYLKIEELSK-VQEDLKNENELHGMV 771

Search completed: March 18, 2002, 09:58:37  
 Job time: 395 sec

..Tue Mar 19 10:57:58 2002

us-09-596-746a-52.rapn

Page 9

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:49 ; Search time 68.77 Seconds

(without alignments)  
620.296 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750

Sequence: 1 MACNNSGKQGNSTANSANADES.....KAAQAAALTNVKEITNPVA 560

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR\_68:\*

1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	33.7	210	2	G70218
2	921	33.5	210	2	S69927
3	886	32.2	212	2	S20543
4	881	32.0	207	2	S37727
5	878	31.9	207	2	S69919
6	877	31.9	212	2	S69921
7	874	31.8	200	2	I40122
8	853	31.0	177	2	I40129
9	846	30.8	179	2	S54187
10	844	30.7	179	2	S54188
11	834.5	30.3	180	2	S54191
12	832	30.3	179	2	S54186
13	811	29.5	189	2	S70263
14	798	29.0	207	2	I40271
15	794	28.9	200	2	S54198
16	786	28.6	207	2	I40276
17	759.5	27.6	199	2	S54197
18	757.5	27.5	211	2	I40145
19	756.5	27.5	193	2	S70279
20	700	25.5	210	2	I40144
21	696.5	25.3	193	2	S70287
22	683.5	24.9	211	2	I40277
23	676.5	24.6	209	2	S69918
24	676.5	24.6	201	2	I40273
25	674.5	24.5	191	2	S70278
26	674.5	24.5	191	2	I40153
27	673.5	24.5	211	2	I40268
28	672.5	24.5	193	2	S70280
29	672	24.4	212	2	S69922

#### ALIGNMENTS

30	667	24.3	194	2	S70268	outer surface prot
31	666.5	24.2	193	2	S70265	outer surface prot
32	666.5	24.2	211	2	I40278	outer surface prot
33	666	24.2	194	2	S70277	outer surface prot
34	665.5	24.2	209	2	S69917	outer surface prot
35	665	24.2	190	2	S70273	outer surface prot
36	661	24.0	212	2	S70254	outer surface prot
37	657.5	23.9	193	2	S70276	outer surface prot
38	656.5	23.9	193	2	S70274	outer surface prot
39	653.5	23.8	180	2	S54189	outer surface prot
40	653.5	23.8	191	2	S70284	outer surface prot
41	652.5	23.7	211	2	S69932	outer surface prot
42	646	23.5	212	2	I40279	outer surface prot
43	642.5	23.4	209	2	I40142	outer surface prot
44	635	23.1	212	2	I40143	outer surface prot
45	634.5	23.1	203	2	I40108	outer surface prot

RESULT: 1  
G70218:  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence.revision 13-Feb-1998 #ext\_change 21-Jul-2000  
C:Accession: G70218; I40269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: G70218  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <KLE>  
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu  
A:Reference number: I40269; MUID:96025162  
A:Accession: I40269  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49497; NID:g707092; PIDN:BA008457.1; PID:g765684  
R:Jauris-Helpe, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.  
Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os  
A:Reference number: S37726; MUID:93268136  
A:Accession: S37726  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <JAD>  
A:Cross-references: EMBL:x69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:Rilvey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di  
A:Reference number: S70235; MUID:96296448  
A:Accession: S70281  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:I42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212  
A:Experimental source: strain IP2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 33.7%; Score 928; DB 2; Length 210;  
Best Local Similarity 98.4%; Pred. No. 1.6e-34;  
Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTSANSADSEYKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 60  
Db 17 ISCNNSKRDGTSANSADSEYKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 76  
QY 61 IGGKIHONNGIDTEYNNHNGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 120  
Db 77 IGGKIHONNGIDTEYNNHNGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 136  
QY 121 KLEKHTDLCRGEVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSRAKEMLANSVK 180  
Db 137 KLEKHTDLCRGEVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSRAKEMLANSVK 196  
QY 181 ELTSPVVAESP 191  
Db 197 ELTSPVVAESP 207

## RESULT 2

S69927  
outer surface protein C precursor - Lyme disease spirochete (strain PKA)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
A:Variety: strain PKA  
C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69927; S72669  
R:Jauris-Heipke, S.; Liegli, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek, E.  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia  
A:Reference number: 140047; MUID:95395018  
A:Accession: S69927  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69589  
A:Experimental source: strain PKA  
R:Jauris, S.  
submitted to the EMBL Data Library, February 1994  
A:Reference number: S72669  
A:Accession: S72669  
A:Molecule type: DNA  
A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAM>  
A:Cross-references: EMBL:X69589  
C:Genetics:  
A:Gene: ospC  
superfamily: Lyme disease spirochete surface protein C

Query Match 33.5%; Score 921; DB 2; Length 210;  
Best Local Similarity 97.4%; Pred. No. 3.2e-34;  
Matches 186; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSKRDGTSANSADSEYKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 60  
Db 17 ISCNNSKRDGTSANSADSEYKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 76  
QY 61 IGGKIHONNGIDTEYNNHNGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 120  
Db 77 IGGKIHONNGIDTEYNNHNGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 136  
QY 121 KLEKHTDLCRGEVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSRAKEMLANSVK 180  
Db 137 KLEKHTDLCRGEVTDADAKAELIKTNGTKGAEEIGKLFESVEVLSRAKEMLANSVK 196  
QY 181 ELTSPVVAESP 191  
Db 197 ELTSPVVAESP 207

## RESULT 3

S20543

S37727

outer surface protein C precursor - Lyme disease spirochete  
M:Alternate names: membrane-associated protein PC  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jul-2000  
C:Accession: S20543; I40107; S69931; S70275; S34175  
R:Fuchs, R.; Jauris, S.; Lottspeich, F.; Preac-Mursic, V.; Wilske, B.; Soutschek, E.  
Mol. Microbiol. 6, 503-509, 1992  
A:Title: Molecular analysis and expression of a Borrelia burgdorferi gene encoding a  
A:Reference number: S20543; MUID:92219995  
A:Accession: S20543  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <FUC>  
A:Cross-references: EMBL:X62162; NID:939392; PIDN:CAA44093.1; PID:939393  
R:Thiesen, M.; Frederiksen, B.; Lebeck, A.M.; Vuust, J.; Hansen, K.  
J. Clin. Microbiol. 31, 2570-2576, 1993  
A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os  
A:Reference number: I40105; MUID:94075528  
A:Accession: I40107  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-205 <RES>  
A:Cross-references: EMBL:X73624; NID:9313273; PIDN:CAA52003.1; PID:9313274  
A:Experimental source: strain DK26  
R:Jauris-Heipke, S.; Liegli, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69931  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-212 <JAU>  
A:Cross-references: EMBL:X69590  
A:Experimental source: strain Fwudi  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1994  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in ospC variation in Lyme di  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70275  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-212 <LIV>  
A:Cross-references: EMBL:L42883; NID:9858709; PIDN:AAB36991.1; PID:91695209  
A:Experimental source: strain JSB  
C:Genetics:  
A:Gene: ospC  
superfamily: Lyme disease spirochete surface protein C

Query Match 32.2%; Score 886; DB 2; Length 212;  
Best Local Similarity 90.7%; Pred. No. 1.1e-32;  
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

QY 195 SMSGKGDASSTNPDESAGKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 254  
Db 20 NNSGKGDASSTNPDESAGKGNLTETISKKTIDTSNAVLAVKEVALLSIDETIAKA 79  
QY 255 OKIDNNNGIAALNNONGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 314  
Db 80 OKIDNNNGIAALNNONGSLAGAVAIISLTIKOKLDGLKNEGLKEKIDAKKCEFTFN 139  
QY 315 LKSGHADLGKODATDDHAKAALIKTHATTDGKAEKDFESVEGLKAAQVALTNSVKE 374  
Db 140 LKSGHADLGKODATDDHAKAALIKTHATTDGKAEKDFESVEGLKAAQVALTNSVKE 199  
QY 375 LGHRNNSGDSASTNP--DESAKGP 397  
Db 200 L-----TSPVVAESP 212

## RESULT 4

S37727

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-May-2000

C:Accession: S37727

R:Jauris-Helpeke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W  
Med. Microbiol. Immunol. 182, 37-50, 1993

A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)

A:Reference number: S37726; MUID:93268136

A:Accession: S37727

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-207 <JAU>

A:Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA5305.1; PID:g311394

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 32.0%; Score 881; DB 2; Length 207;

Best Local Similarity 99.5%; Pred. No. 1,8e-32;

Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNDESAKGNLTVISKITDTSNAFLAVKEVALLSSIDELSKAIGKTK 438

DB 20 NNSGDSASTNDESAKGNLTVISKITDTSNAFLAVKEVALLSSIDELSKAIGKTK 79

QY 439 NNGTLDNEANRRESLIAGAYEISKLTITOKLSVLNSELKKEKAKDCOKFTTKLDKSH 498

DB 80 NNGTLDNEANRRESLIAGAYEISKLTITOKLSVLNSELKKEKAKDCOKFTTKLDKSH 139

QY 499 AELGIQSVODDNNAKKAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSKELTNV 558

DB 140 AELGIQSVODDNNAKKAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSKELTNV 199

QY 559 VA 560

DB 200 VA 201

RESULT 5

S69919

outer surface protein C precursor - Borrelia garinii (strain PTrob)

C:Species: Borrelia garinii

A:Variety: strain PTrob

C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000

C:Accession: S69919

R:Jauris-Helpeke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia

A:Reference number: 140047; MUID:95395018

A:Accession: S69919

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-207 <JAU>

A:Cross-references: EMBL:X83554; NID:g872027; PIDN:CAA5854.1; PID:g872028

A:Experimental source: strain PTrob

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

F.1-18/Domain: signal sequence #status predicted <SIG>

F.19-207/Product: outer surface protein C #status predicted <MAT>

Query Match 31.9%; Score 878; DB 2; Length 207;

Best Local Similarity 98.9%; Pred. No. 2,4e-32;

Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNDESAKGNLTVISKITDTSNAFLAVKEVALLSSIDELSKAIGKTK 438

DB 20 NNSGDSASTNDESAKGNLTVISKITDTSNAFLAVKEVALLSSIDELSKAIGKTK 79

QY 439 NNGTLDNEANRRESLIAGAYEISKLTITOKLSVLNSELKKEKAKDCOKFTTKLDKSH 498

DB 80 NNGTLDNEANRRESLIAGAYEISKLTITOKLSVLNSELKKEKAKDCOKFTTKLDKSH 139

QY 499 AELGIQSVODDNNAKKAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSKELTNV 558

DB 140 AELGIQSVODDNNAKKAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSKELTNV 199

QY 559 VA 560

DB 200 VA 201

RESULT 6

S69921

outer surface protein C precursor - Borrelia afzelii (strain PBO)

C:Species: Borrelia afzelii

A:Variety: strain PBO

C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000

C:Accession: S69921; 140047

R:Jauris-Helpeke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre

A:Reference number: 140047; MUID:95395018

A:Accession: S69921

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-212 <JAU>

A:Cross-references: EMBL:X81521; NID:g804954; PIDN:CAA57241.1; PID:g804955

A:Experimental source: strain PBO

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 31.9%; Score 877; DB 2; Length 212;

Best Local Similarity 89.8%; Pred. No. 2,7e-32;

Matches 184; Conservative 4; Mismatches 3; Indels 14; Gaps 2;

QY 195 NNSGKGDASTNPADSAKGNLTVISKITDTSNAFLAVKEVALLSSIDELAKKAG 254

DB 20 NNSGKGDASTNPADSAKGNLTVISKITDTSNAFLAVKEVALLSSIDELAKKAG 79

QY 255 OKIDNNNGLAALNONGSLAGAYASTLTETKLSKLNLEELKTEIAKAKCSEFTNK 314

DB 80 OKIDNNNGLAALNONGSLAGAYASTLTETKLSKLNLEELKTEIAKAKCSEFTNK 139

QY 315 LKSGHADLCKODATDHAHAAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSK 374

DB 140 LKSGHADLCKODATDHAHAAILTKHTGDKGAKLELEFKLSLSLSKAAQALTNVSK 199

QY 375 LCHRNNSGDSASTNP--DESAKGP 397

DB 200 L-----TSPVAESPCKP 212

RESULT 7

I40122

outer surface protein C - Lyme disease spirochete (strain DK6) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000

C:Accession: I40122; S34177

R:Thielsen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.

J. Clin. Microbiol. 31, 2570-2576, 1993

A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os

A:Reference number: I40105; MUID:94075528

A:Accession: I40122

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-200 <RES>

A:Cross-references: EMBL:X73626; NID:g313277; PIDN:CAA52005.1; PID:g313278

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match 31.8%; Score 874; DB 2; Length 200;  
Best Local Similarity 98.9%; Pred. No. 3.5e-32;  
Matches 179; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNPDESAGPNTLEISKITDSNAFLAVKEVEALLSIDELSKAIGKKIK 438  
|||||  
DB 20 NNSGDSASTNPDESAGPNTLEISKITDSNAFLAVKEVEALLSIDELSKAIGKKIK 79  
QY 439 NGCTLDENNRNRESILAGAYEISKLITOKLSVLSNSELKKTREKAKDCSQKFTTKLKDSH 498  
|||||  
DB 80 NGCTLDENNRNRESILAGAYEISKLITOKLSVLSNSELKKTREKAKDCSEKFTTKLKDSH 139  
QY 499 AELIGISVODDAAKKAIIKTHGKTGKAGKELEFKSLESLSKAQAALTNVSKELTNV 558  
|||||  
DB 140 AELIGISVODDAAKKAIIKTHGKTGKAGKELEFKSLESLSKAQAALTNVSKELTNV 199  
559 V 559  
DB 200 V 200

RESULT 8  
140129  
outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I40129; S54199  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40129  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-177 <RES>  
A:Cross-references: EMBL:X84783; NID:9793825; PIDN:CAA59254.1; PID:9793826  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 31.0%; Score 853; DB 2; Length 177;  
Best Local Similarity 98.9%; Pred. No. 2.5e-31;  
Matches 175; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 11 NTSANSADSESVKGNLTETISKRITDSNAVLAVKEVEALLSIDELSKAIGKKIHONG 70  
|||||  
DB 1 NTSANSADSESVKGNLTETISKRITDSNAVLAVKEVEALLSIDELSKAIGKKIHONG 60  
QY 71 LDTEYNHNSILAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTTKLKEKHTDLG 130  
|||||  
DB 61 LDTEYNHNSILAGAYASTLIKOKLDGKNEGLKEKIDAKKCEFTTKLKEKHTDLG 120  
QY 131 KEVYTDADAEALIKTNGTKGAEELGKLFESVEVLSKAKEMLANVSEKELSPV 187  
|||||  
DB 121 KRGVTDADAEALIKTNGTKGAEELGKLFESVEVLSKAKEMLANVSEKELSPV 177

RESULT 9  
S54187  
outer surface protein C - Borrelia afzelii (strain DK2) (fragment)  
C:Species: Borrelia afzelii  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
C:Accession: I40115; S54187  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein Ospc.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40115  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-179 <THE>

A:Cross-references: EMBL:X84766; NID:9793801; PIDN:CAA59237.1; PID:9793802  
A:Experimental source: strain DK2  
A:Note: submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 30.8%; Score 846; DB 2; Length 179;  
Best Local Similarity 99.4%; Pred. No. 5.1e-31;  
Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 202 DSASTNPDESAGPNTLEISKITDSNAFLAVKEVEVLVLSIDELAKKAIGKIDNNN 261  
|||||  
DB 1 DSASTNPDESAGPNTLEISKITDSNAFLAVKEVEVLVLSIDELAKKAIGKIDNNN 60  
QY 262 GLAALNNONGSLAGAYASTLIKTEKLSKLNLEELKTEIAKAKKCEFTTKLKSCHAD 321  
|||||  
DB 61 GLAALNNONGSLAGAYASTLIKTEKLSKLNLEELKTEIAKAKKCEFTTKLKSCHAD 120  
QY 322 LGKODATDHAARAIKTHATTGKAKEFKDLFESVEGLKAAQVALTNVSKEL 375  
|||||  
DB 121 LGKODATDHAARAIKTHATTGKAKEFKDLFESVEGLKAAQVALTNVSKEL 174

RESULT 10  
S54188  
outer surface protein C - Borrelia afzelii (strain DK3) (fragment)  
C:Species: Borrelia afzelii  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
C:Accession: I40116; S54188  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein Ospc.  
A:Reference number: I40104; MUID:95286481  
A:Accession: I40116  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-179 <THE>  
A:Cross-references: EMBL:X84771; NID:9793803; PIDN:CAA59242.1; PID:9793804  
A:Experimental source: strain DK3  
A:Note: submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 30.7%; Score 844; DB 2; Length 179;  
Best Local Similarity 99.4%; Pred. No. 6.3e-31;  
Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 202 DSASTNPDESAGPNTLEISKITDSNAFLAVKEVEVLVLSIDELAKKAIGKIDNNN 261  
|||||  
DB 1 DSASTNPDESAGPNTLEISKITDSNAFLAVKEVEVLVLSIDELAKKAIGKIDNNN 60  
QY 262 GLAALNNONGSLAGAYASTLIKTEKLSKLNLEELKTEIAKAKKCEFTTKLKSCHAD 321  
|||||  
DB 61 GLAALNNONGSLAGAYASTLIKTEKLSKLNLEELKTEIAKAKKCEFTTKLKSCHAD 120  
QY 322 LGKODATDHAARAIKTHATTGKAKEFKDLFESVEGLKAAQVALTNVSKEL 375  
|||||  
DB 121 LGKODATDHAARAIKTHATTGKAKEFKDLFESVEGLKAAQVALTNVSKEL 174

RESULT 11  
S54191  
outer surface protein C - Borrelia afzelii (strain DK8) (fragment)  
C:Species: Borrelia afzelii  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
C:Accession: I40119; S54191  
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein Ospc.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:57 ; Search time 39.62 Seconds

(without alignments)  
518.231 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750

Sequence: 1 MACNNGKNGNTSANSADSE.....KAAQAALTNVKELTNPVVA 560

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	33.7	210	1	OSCL_BORBU
2	886	32.2	212	1	OSCL_BORBU
3	402.5	14.6	214	1	VM24_BORHE
4	398	14.5	215	1	VM03_BORHE
5	210.5	7.7	1957	1	Y086_SCHPO
6	207.5	7.5	1790	1	US01_YEAST
7	207.5	7.5	2869	1	RBP1_PLAVB
8	204.5	7.4	1509	1	MYSN_ACACA
9	193.5	7.0	2017	1	MYSN_DROME
10	193.5	7.0	2116	1	MYS2_DICDI
11	193	7.0	1433	1	REST_CHICK
12	192	7.0	1966	1	MYSB_CAEL
13	189	6.9	1022	1	SCA4_RICPR
14	189	6.9	1427	1	REST_HUMAN
15	187	6.8	875	1	ZIP1_YEAST
16	187	6.8	1164	1	BAC_STRAG
17	186	6.8	1167	1	CAGA_HELPJ
18	185	6.7	1189	1	SCIT_CHICK
19	185	6.7	1251	1	RBP2_PLAVB
20	184.5	6.7	1636	1	BUD3_YEAST
21	183.5	6.7	1189	1	YJH6_YEAST
22	183.5	6.7	1940	1	MYH3_CHICK
23	182.5	6.6	1786	1	LMB1_MOUSE
24	180.5	6.6	775	1	YHGE_BACSU
25	180	6.5	3672	1	LML2_CAEL
26	179.5	6.5	2411	1	MYS4_DROME
27	179	6.5	1940	1	MYH3_RAT
28	178.5	6.5	727	1	MPL1_ARATH
29	178.5	6.5	1290	1	XCP3_XENLA
30	177.5	6.5	1786	1	LMB1_HUMAN
31	177	6.4	1875	1	MLP1_YEAST
32	177	6.4	2748	1	NUM1_YEAST
33	176.5	6.4	998	1	SCA4_RICAK

## ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD:	PRT:	210 AA.
AC	007337	OSCL_BORBU			
DT	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB19.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Plasmodium falciparum.				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 35210 / B31;				
RA	MEDLINE=93268136; PubMed=8098841;				
RA	Jauris-Helpke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,				
RT	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 35210 / B31;				
RA	MEDLINE=9339332; PubMed=8478108;				
RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant				
RT	major outer surface protein of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 35210 / B31;				
RA	MEDLINE=94041630; PubMed=8225587;				
RA	Padula S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 35210 / B31;				
RA	MEDLINE=96025162; PubMed=7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.";				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 35210 / B31;				
RA	MEDLINE=98065943; PubMed=9403685;				
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Laughlin R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

P02564 rattus norv  
O27991 bos taurus  
P12845 caenorhabd  
O05000 podocoryne  
P12883 homo sapien  
P27895 saccharomyc  
O9m714 nicotiana t  
P23745 plasmodium  
P13539 mesocricetu  
O08696 drosophila  
P79293 sus scrofa  
P35749 homo sapien

RA Uterback T., Matthey L., McDonald L., Attiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X69596; CAA49306.1; -;  
 CC EMBL; U01894; AAA16058.1; -;  
 CC EMBL; D49497; BAA08457.1; -;  
 CC EMBL; AE000792; AAC66329.1; -;  
 CC TIGR; BB819; -;  
 CC InterPro: IPR001800; Lipoprotein\_6.  
 CC Pfam; PF01441; Lipoprotein\_6; 1.  
 CC ProDom; PD001149; Lipoprotein\_6; 1.  
 CC PROSITE; PS00013; PROKAR.LIPOPROTEIN.1.  
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 CC Complete proteome.  
 CC SIGNAL  
 CC CHAIN 1 18 BY SIMILARITY  
 CC FT LIPID 19 210 OUTER SURFACE PROTEIN C.  
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 CC SQ SEQUENCE 210 AA; 22340 MW; 744FC978F91777BF CRC64;  
 CC -----  
 CC Query Match 33.7%; Score 928; DB 1; Length 210;  
 CC Best Local Similarity 98.4%; Pred. No. 5e-33;  
 CC Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MACNNSGKDGNTSANSADSVKCPNLTEISKRTDSNAVLAVKEVELLSSIDEIAKA 60  
 CC DB 17 ISCNNSKDGNTSANSADSVKCPNLTEISKRTDSNAVLAVKEVELLSSIDEIAKA 76  
 CC QY 61 IGGKTIHONNGIDTYENHNGSLAGAVAI STLTKKLDGKNEGKEKRTDAKKCSEFTTN 120  
 CC DB 77 IGGKTIHONNGIDTYENHNGSLAGAVAI STLTKKLDGKNEGKEKRTDAKKCSEFTTN 136  
 CC DB 121 KIKKHHDTLGKGGVTDADAKAAILTKGTGKGAEEIGKLFESVEVLSKAKEMLANSVK 180  
 CC DB 137 KIKKHHDTLGKGGVTDADAKAAILTKGTGKGAEEIGKLFESVEVLSKAKEMLANSVK 196  
 CC QY 181 ELTSPVVAESP 191  
 CC DB 197 ELTSPVVAESP 207  
 CC  
 CC RESULT 2  
 CC OSC2\_BORBU STANDARD; PRT; 212 AA.  
 CC ID OSC2\_BORBU 008137;  
 CC AC 008137;  
 CC DT 15-DEC-1998 (Rel. 37, Created)  
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 CC DE OUTER SURFACE PROTEIN C PRECURSOR (PC).  
 CC GN OSPC.  
 CC OS Borrelia burgdorferi (Lyme disease spirochete).  
 CC OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 CC OX NCBI\_TaxID=139;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 CC RC STRAIN=PKO;  
 CC RX MEDLINE=9221995; PubMed=1560779;  
 CC RA Fuchs R., Jauris S., Lotspeich F., Preac-Mursic V., Wilske B.,

RA Soutschek E.;  
 RT "Molecular analysis and expression of a Borrelia burgdorferi gene  
 RT encoding a 22 kDa protein (pc) in Escherichia coli.";  
 RL Mol. Microbiol. 6:503-509(1992).  
 CC [2]  
 CC SEQUENCE OF 1-205 FROM N.A.  
 CC RC STRAIN=DK26;  
 CC RX MEDLINE=94075528; PubMed=8253951.  
 CC RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
 CC "Polymorphism in ospC gene of Borrelia burgdorferi and  
 CC immunoreactivity of ospC protein: implications for taxonomy and for  
 CC use of ospC protein as a diagnostic antigen.";  
 CC J. Clin. Microbiol. 31:2570-2576(1993).  
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 CC ANCHOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X62162; CAA44093.1; -;  
 CC EMBL; X73624; CAA52003.1; -;  
 CC InterPro: IPR001800; Lipoprotein\_6.  
 CC Pfam; PF01441; Lipoprotein\_6; 1.  
 CC ProDom; PD001149; Lipoprotein\_6; 1.  
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;  
 CC Complete proteome.  
 CC SIGNAL  
 CC CHAIN 1 18 BY SIMILARITY  
 CC FT LIPID 19 212 OUTER SURFACE PROTEIN C.  
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 CC SQ SEQUENCE 212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;  
 CC -----  
 CC Query Match 32.2%; Score 886; DB 1; Length 212;  
 CC Best Local Similarity 90.7%; Pred. No. 2.9e-31;  
 CC Matches 166; Conservative 2; Mismatches 3; Indels 14; Gaps 2;  
 CC  
 CC QY 195 SNSGKGDSDASTNPADDSAGPNLTETSKRTDSNAVLAVKEVELLSSIDEIAKAIG 254  
 CC DB 20 NNSGKGDSDASTNPADDSAGPNLTETSKRTDSNAVLAVKEVELLSSIDEIAKAIG 79  
 CC QY 255 OKIDNNNGLAALNNONGSLAGAVAI STLTKKLSKKNLEKTELAKKKCSSEFTNK 314  
 CC DB 80 OKIDNNNGLAALNNONGSLAGAVAI STLTKKLSKKNLEKTELAKKKCSSEFTNK 139  
 CC QY 315 LKSGHADLGKQDATTDDAKAAILTKHTATDKAKEKFDLFESVEGLLKAQVALTSYKE 374  
 CC DB 140 LKSGHADLGKQDATTDDAKAAILTKHTATDKAKEKFDLFESVEGLLKAQVALTSYKE 199  
 CC QY 375 LGHRNNSGDSASTNP-DESAKGP 397  
 CC DB 200 L-----TSPVVAESP 212  
 CC  
 CC RESULT 3  
 CC VM24\_BORHE STANDARD; PRT; 214 AA.  
 CC ID VM24\_BORHE  
 CC AC P32778;  
 CC DT 01-OCT-1993 (Rel. 27, Created)  
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 CC DE 01-FEB-1994 (Rel. 28, Last annotation update)  
 CC DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.  
 CC GN VMP24.  
 CC OS Borrelia hermsli.  
 CC OC Plasmid.  
 CC OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 CC OX NCBI\_TaxID=140;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.

RC STRAIN-SSP, HSI SEROTYPE 24;  
 RA MEDLINE-93133110; PubMed-1484486;  
 RT Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 "Subtelomeric expression regions of Borrelia hermsli linear plasmids  
 are highly polymorphic.";  
 RL Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, L04786; AAA22964.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 DR Outer membrane; Lipoprotein; Signal; Plasmid.  
 FT SIGNAL 1 18 PROBABLE MAJOR OUTER MEMBRANE  
 FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT SEQUENCE 214 AA: 22541 MW: F1583F510246F7C7 CRC64;

Query Match 14.6%; Score 402.5; DB 1; Length 214;  
 Best Local Similarity 44.9%; Pred. No. 4.9e-11;  
 Matches 92; Conservative 37; Mismatches 59; Indels 17; Gaps 5;

OY 1 MACNNSGKNGTNSANSADSVKGP---NLTEISKRTTDSNVLAVKEVALLSIDEI 56  
 DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKVSKIKKASFAASVKEVTLVKSVDL 71  
 OY 57 AKAIGKRIHONGDTEYNHNGSLAGAYASTLIKOKL-----DGLKNGLEKIDA 110  
 DB 72 -AKAIGKRIKNDGSLDTEAGNGSLIAGVSVAVKIKVGLLETTSGISNE-LTKRIIE 129  
 OY 111 AKKCEFTFNKLEKHTDGLGEGVTDADAKKAILKTNGTKTGAEBLGFESVEVLSKA 170  
 DB 130 VSKSAEAFYTKSKHTDGLGEGVTDADAKKAILKTNGTKTGAEBLGFESVEVLSKA 189  
 OY 171 AKEMLANSVKELTSPVAVSPAMGS 195  
 DB 190 AGEVEBAIKELTAPVAKKAPSONN 214

RESULT 4  
 ID VMO3 BORHE STANDARD: PRT: 215 AA.  
 AC 002448: 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.  
 GN VMP3.  
 OS Borrelia hermsli.  
 OC Plasmid.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=140;  
 RN NCB1  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SSP, HSI SEROTYPE 3;  
 RA MEDLINE-93133110; PubMed-1484486;  
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;  
 RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

RT are highly polymorphic.";  
 RL Mol. Microbiol. 6:3299-3311(1992).  
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
 FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
 ANCHOR.  
 CC -1- SIMILARITY: STRONG, TO VMP24.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, L04789; AAA22967.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 DR Outer membrane; Lipoprotein; Signal; Plasmid.  
 FT SIGNAL 1 18 PROBABLE MAJOR OUTER MEMBRANE  
 FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE  
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT SEQUENCE 215 AA: 23139 MW: 684C7AD35F87C771 CRC64;

Query Match 14.5%; Score 398; DB 1; Length 215;  
 Best Local Similarity 45.4%; Pred. No. 7.5e-11;  
 Matches 93; Conservative 35; Mismatches 59; Indels 16; Gaps 5;

OY 1 MACNNSGKNGTNSANSADSVKGP---NLTEISKRTTDSNVLAVKEVALLSIDEI 56  
 DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKVSKIKKASFAASVKEVTLVKSVDL 71  
 OY 57 AKAIGKRI-HONGDTEYNHNGSLAGAYASTLIKOKL-----NEGKRIKIDA 110  
 DB 72 -AKAIGKRIKNDNSNFEDNDHNGSLIAGVFOYILTVKAKLISLEDTIGISDLKTEVGM 130  
 DB 131 VKKESAEFTQVKSHTDGLGEGVTDADAKKAILKTNGTKTGAEBLGFESVEVLSKA 190  
 OY 111 AKKCEFTFNKLEKHTDGLGEGVTDADAKKAILKTNGTKTGAEBLGFESVEVLSKA 170  
 DB 171 AKEMLANSVKELTSPVAVSPAMGS 195  
 DB 191 ANDAVETVIKELTASVAKKAPSONN 215

RESULT 5  
 ID YD86\_SCHPO STANDARD: PRT: 1957 AA.  
 AC 010411: 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.  
 GN SPAC1F3.06C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN NCB1  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 270690; CAA4624.1; -  
KW Hypothetical protein.  
SQ Sequence 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 7.7%; Score 210.5; DB 1; Length 1957;  
Best Local Similarity 22.4%; Pred. No. 0.065;  
Matches 157; Conservative 116; Mismatches 218; Indels 209; Gaps 35;

```

OY 8 KGGNTSANSADK---SVKGNLTETSKITDTSNAVLAV--KVEALLSIDELAKAIGK 63
DB 360 KDSRTNSOLEEEMVELKSNRT--HSQJLTDSEKLSSEQENKSLKSGSIDE----- 410
DB 64 KIHONN-----GLDTEYHNHNSLAGAYASTLTKOKLIDGKNEGLK--EKIDA-A 111
DB 411 --YNNLSSKDKMVKVQSSQLEEARSLAHATGKLAIEINSEED--FONKKIKDFEIQDL 467
OY 112 KRCSEFTFKLKEKHDTDLCEGVTDADAKAELTKNGTKTKAEGLKLFESV--EVLISK 169
DB 468 RAGLWSSSSEKELKSLIKKQDLNLEPEQIKK---OKKVESTQSSLSLQDILINE 523
OY 170 AAK-EMLANSVKELTSPVVAESPAMGSGSGASTNPADSEAKGNLTETSK--- 224
DB 524 KKKHEYESQELNELKGELOTEI---SNSEHLSQSLTLAEKEAAVATNNELSSKSL 579
OY 225 IYDSNAF-----VLAVKEVETLVLSIDELAKK--AIGKIDNNNGI----- 263
DB 580 QTLCAFOBKLAQSVWQLEKNEONFSSLDTSFKKLNESHQELNNHQTITQOLKDTSSKL 639
OY 264 -----AALNNGSLAGAYASTLTKETSKL----- 291
DB 640 QOLLEKLANFEQESTLSDENNLD-----RTLLKLESNKSLIKQEDVDLSLE 688
OY 292 KNELEKLEIAKAK-----KCESEFTNKLKSGHADLGKO-----DATDDHAKAA 335
DB 689 KNIQTLKEDLKRSEALRFSKLEAKNLEVIDNKGKHETLEAQRNDLHSLSDAKNTNA 748
OY 336 ILKTHATDCKAKEKDFEFSEVEGL-----LKAQVALTNSVVELGH----- 377
DB 749 ILSEELT--KSEDEYKLTANVELLTQDSKAMKOSFTSLVNSYSISNLVHELDDHVM 806
OY 378 --RNNS--GGDSASTNPDESAKGNLTYSKIKTDSNAFLAVKEV--EALLSIDELSK 431
DB 807 QSONNTLSESEKLTDCENLTQONMTLI-----DNVQKLMHKNHVNQESVSEKLEVNG 860
OY 432 AIGKKIKK--DGLTNEANRNSLAGAVEISK---LITOKLSVNS----- 473
DB 861 KLSLDLKNLSSLVNVAISDNQJLTQLAELSKNYSLEQSAQSLKSLAEKQLLHT 920
OY 474 --EEL-----KKRIKEAK--DCSQKFTTKLKDASHAELGQSVODDD--AKKALIKT 518
DB 921 ENEELHRLDKLTGKLTIESKSSDLGKTLTAQOE-----ISNLKEEMASOALITSV 974
OY 519 HGTQDKAKLELELFKSLSEKRAQAALTNVSKVELTNPV 558
DB 975 KSKLD-----ETLSKSSK--LEADIEHLKNKV 999

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae."
RN J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RN Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RN Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Query Match 7.5%; Score 207.5; DB 1; Length 1790;  
Best Local Similarity 21.4%; Pred. No. 0.079;  
Matches 143; Conservative 104; Mismatches 259; Indels 161; Gaps 22;

```

OY 3 CNGSKDG-----NTSANSADSEKGNLTETSKITDTSNAVLAVKEVALLS 52
DB 929 CNNLSKEKEHISKELVEYKSRGSHDNLV--AKLEKAKSLANNYKMQA--EKESLIKA 984
OY 53 IDELAKAIGKKIHONNGLD--TEYNHNSLAGAY-----AISTL-----IKOK 95

```



OS Acanthamoeba castellanii (Acanthamoeba).  
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
 RX NCBI\_TaxID=5755;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=87308395; PubMed=3040773;  
 RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;  
 RT "Complete nucleotide sequence and deduced polypeptide sequence of a  
 RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a  
 RT hinge in the rodlike tail."  
 RL J. Cell Biol. 105:913-925(1987).  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC). 2 ALKALI  
 CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MLC-2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS  
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE  
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.  
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING  
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: Y00624; CA68663.1; -  
 CC PIR: A27224; A27224.  
 DR HSSP; P08799; IMND.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 CC Myosin: Coiled coil; Actin-binding; Calmodulin-binding;  
 CC Methylation; Alkylation; Phosphorylation; Multigene family.  
 FT DOMAIN 1 789 MYOSIN HEAD-LIKE.  
 FT DOMAIN 790 819 IQ.  
 FT DOMAIN 848 1509 COILED COIL (POTENTIAL).  
 FT DOMAIN 848 1226 ALPHA-HELICAL TAILPIECE (S2).  
 FT DOMAIN 1227 1252 HINGE.  
 FT DOMAIN 1253 1509 LIGHT MEROMYOSIN (LMM).  
 FT DOMAIN 1253 1482 ALPHA-HELICAL TAILPIECE (LMM).  
 FT DOMAIN 1483 1509 NONHELICAL TAILPIECE.  
 FT NP\_BIND 182 189 ATP.  
 FT DOMAIN 660 682 ACTIN-BINDING.  
 FT DOMAIN 766 780 ACTIN-BINDING.  
 FT MOD\_RES 133 133 METHYLATION (TRI-) (POTENTIAL).  
 FT MOD\_RES 700 700 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 1489 1489 PHOSPHORYLATION.  
 FT MOD\_RES 1494 1494 PHOSPHORYLATION.  
 FT MOD\_RES 1499 1499 PHOSPHORYLATION.  
 FT SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 7.4%; Score 204.5; DB 1; Length 1509;  
 Best Local Similarity 22.2%; Pred. No. 0.086;  
 Matches 127; Conservative 107; Mismatches 218; Indels 119; Gaps 25;  
 QY 28 EISKITDSNAVLAVKEVEA-----LLSSIDE-----IAAKAIQKIHQNNGLDREYNNH 77

DB 934 ELQETSTSNILQCKRLLEKKGELKASLEEEERNKRALQEAATTKYEESENNELQDKED 993  
 QY 78 NSLLAGAYASTLLIKOKLIDGKNEGKLEKIDA--AKCSEFTNKLKEKHTLDGREGV 134  
 DB 994 EA-----AAHDSIKKKEEDLSRE-LRETQDALADENISETILRSKL--KNTERGADDV 1043  
 QY 135 TDA--DAKEALLTKNGTKKAEEELKLFEEVEYLSRAKEMLANSKVELTSPVVAESPA 192  
 DB 1044 RNELDVATYATKLOLEKTKLSLEELAQTRALEE-ENSGKSAASAKQLQOQLEDAESE 1102  
 QY 193 MGSNGSGGDSASTNPDESAGPNLITEISKITDSNAFVLAWEFTVLSIDELAKKA 252  
 DB 1103 VDSLKSK--LSAAEKSLTKAKDON-RDLQDQLDE-----RTVRANVVK-QKA 1147  
 QY 253 IGQKI-DNNGGLAALNNONGSLAGAYASTLITEKLSKLNLELTKETAKAKCSEEF 311  
 DB 1148 LEAKLTLELDQVTLADQCKNAQAATKLKQVDETKRRLLEAASARLEKERNALDE 1207  
 QY 312 TNKLKS---GHADLGKQDAVDHAKKAIL-----KTHATTGKAKKEFDLFEEVGL 360  
 DB 1208 VAQLTADLDERDSGAQQRKRLNTRISELOSELENAPRTGASSEYKRLGELERLEE 1267  
 QY 361 LKAAQVALTNSVKELGHRN-----NSGDSASTNPDESAGPNLTVISKI-TDSNAFL 413  
 DB 1268 LLTQGEARAAAEKMLDKANLELEELRQADDAANDNKLVKD-----NKRKADDEAR 1321  
 QY 414 LAWEVE-----ALLSIDELSKAIGKIKN-DGTLDEAN--RNSLIGAY 458  
 DB 1322 IQLEEDDANSHADSSRRRLAELEELKRVAKETSDKRAQODKANYORENSLKA--- 1378  
 QY 459 EISKLTITQKLSVNSELKKIKKAEKDCSQKFTTKLDSHAELGIGVODNKAALKTK 518  
 DB 1379 -----DRDSIERNNRAE-----RQVYDLAQDLALSRDSEKRA----- 1414  
 QY 519 HGTKDKGAKLELELEK-----SLESLSK 541  
 DB 1415 ---KEKSEVNEANRELKRVYLDREKRSLELSK 1442  
 RESULT 9  
 ID MYSN\_DROME STANDARD; PRT; 2017 AA.  
 AC G99323;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).  
 GN ZIP.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=90349606; PubMed=2117279;  
 RA Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;  
 RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain  
 RT transcript: sequence of the Drosophila nonmuscle myosin heavy-chain  
 RT splicing in the 5' untranslated sequence."  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).  
 CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR  
 CC CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M35012; AAA28713.1; .  
CC PIR: A36014; A36014.  
CC PIR: B36014; B36014.  
CC HSSP: P08799; 1MND.  
CC FLYbase; FBgn0005634; zip.  
CC InterPro: IPR000048; IQ.  
CC InterPro: IPR002928; Myosin\_tail.  
CC InterPro: IPR001609; myosin\_head.  
CC Pfam; PF00612; IQ; 1.  
CC Pfam; PF00063; myosin\_head; 1.  
CC Pfam; PF01576; Myosin\_tail; 1.  
CC PRINTS; PR00193; MYOSINHEAVY.  
CC PRODOM; PD000355; myosin\_head; 1.  
CC SMART; SM00015; IQ; 1.  
CC SMART; SM00242; MYSC; 1.  
CC PROSITE; PS00096; IQ; 1.  
CC Myosin; Alternative splicing; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding.  
CC MYOSIN HEAD-LIKE.  
FT DOMAIN 1 829  
FT DOMAIN 886 859  
FT NP\_BIND 225 232  
FT NP\_BIND 250 260  
FT DOMAIN 682 694  
FT DOMAIN 705 727  
FT DOMAIN 742 758  
FT DOMAIN 1303 2017  
FT DOMAIN 1303 1970  
FT DOMAIN 1971 2017  
FT VAAPLIC 1 45  
SQ SEQUENCE 2017 AA; 232016 MW; 73E3CB02BBAF2528 CRC64;

Query Match 7.0%; Score 193.5; DB 1; Length 2017;  
Best Local Similarity 23.3%; Pred. No. 0.35;  
Matches 131; Conservative 87; Mismatches 181; Indels 163; Gaps 27;

OY 26 LFEISKRTIDSNVLL-----AYKEVEALLSID-----ELAKAIGKRTIHONGDTEYN 76.  
DB 1136 LIQTLLRIDEESATKATAOKAORELESQAEIOEDLEAKARAKAREKRRD----- 1187  
OY 77 HNGSLLAGAYAISTLIKOKLDGLNGLEKEKIDA-----AKKCEFTFKLEKEKHTD 128  
DB 1188 -----LSELELAKNE-LIDSIDTTAAQOELSKREQELATLKKSLEET 1231  
OY 139 LKEGVTDAADAKAELIKTNGTKGAEEELGKLFESVEYLSAKAKEMLANSVKELTSPVVA 188  
DB 1232 VNHCEVL-ADMMH-----KHSQELNSINDOLENKR-AKTVL----- 1266  
OY 189 ESPAMGNSGKGGDSASINPADESAKGPVLTETISKITDSNAFVLAVK--EYETLVLSID 246  
DB 1267 -EKAGTLEAEVADLATELRVNSROEN--DRRRQASQAEIQLVQAEIERASEQ 1323  
OY 247 ELAKKAIGKIDNNNGMALNNO--NGSLLAGAYAISTLITELKSLKLELEKTIKA 304  
DB 1324 EKCCTK-LQOEAEEN-----ITNQLAEELKASA-----AVKSASNNESOLTER 1364  
OY 305 KKCSEFT-----NKLKSGHADLGKODATDHAH-----AALIKTHATTDKGA 347  
DB 1366 QQLLEETPQKLGSLSKLQIGSEKFAIQEQLEEDDEAKRNERKAETVTTQOELTKKA 1424  
OY 348 KKKFKDLFESVEGLKAAQ--VALINSVKELGHRNNSGDSASTNPDESAGKPNLTIVISK 404  
DB 1425 EEDADLAKLEEGKRRKLNKDIEALEROVKELIAQND-----RLDKSKK-----KIOS 1471  
OY 405 KTTDSNAFLA-----VKEVEALLSIDEL---SKAIGKIKND-GTLDNEARNESLING 456  
DB 1472 ELEDATIELEAORTVLELEKKOKNPNDKTLAEKKAISFOIAOERDTAEAREKE----- 1526

OY 457 AVEISKLTOKSVYNSLEELKKIKKAEANDCSQKFTTKLKDSHAELGICSVODNAKKAL 516  
DB 1527 -----TKVLSTV--SRLEDEAFDKIEDLENKRRT-----LQNELD-----LA 1561  
OY 517 KTHGTRKDGAKELFEKLSLES 538  
DB 1562 NTQGTADKRVHLEKAKRALES 1583  
RESULT 10  
MS2\_DICDI STANDARD; PRT: 2116 AA.  
AC P08799; 1MND.  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.  
GN MHCA.  
OS Dictyostellium discoideum (slime mold).  
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87092266; PubMed=3540939;  
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;  
RT "Conserved protein domains in a myosin heavy chain gene from Dictyostellium discoideum."  
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).  
RN [2]  
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.  
RC STRAIN=AX2;  
RX MEDLINE=90353583; PubMed=2387408;  
RA Lueck-Vietmeter D., Schleicher M., Grabatin B., Wipplert J., Gersch G.;  
RT "Replacement of threonine residues by serine and alanine in a phosphorylatable heavy chain fragment of Dictyostellium myosin II."  
RL FEBS Lett. 269:239-243(1990).  
RN [3]  
RP PHOSPHORYLATION SITES.  
RX MEDLINE=88112226; PubMed=2828113;  
RA Wagle G., Noegel A., Scheel J., Gersch G.;  
RT "Phosphorylation of threonine residues on cloned fragments of the Dictyostellium myosin heavy chain."  
RL FEBS Lett. 227:71-75(1988).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
RX MEDLINE=95345066; PubMed=7619795;  
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Holden H.M., Rayment I.;  
RT "X-ray structures of the myosin motor domain of Dictyostellium discoideum complexed with MgADP, BeFx and MgADP.AlF<sub>4</sub><sup>-</sup>."  
RL Biochemistry 34:8960-8972(1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.  
RX MEDLINE=95345067; PubMed=7619796;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the truncated head of Dictyostellium discoideum myosin to 2.7-A resolution."  
RL Biochemistry 34:8973-8981(1995).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RX MEDLINE=96206189; PubMed=8611530;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the Dictyostellium discoideum myosin motor domain to 1.9-A resolution."  
RL Biochemistry 35:5404-5417(1996).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
RX MEDLINE=97452580; PubMed=9305951;  
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes

RT of the Dictyostelium discoidium myosin motor domain.";  
 RL Biochemistry 36:11619-11628(1997).  
 RA [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RX MEDLINE-98070605; PubMed-9405148;  
 RA Baer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
 RT "X-ray crystal structure and solution fluorescence characterization  
 of Mg<sup>2+</sup>(3')-O-(N-methylanthraniloyl) nucleotides bound to the  
 Dictyostelium discoidium myosin motor domain.";  
 RL J. Mol. Biol. 274:394-407(1997).  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN.  
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (HMC), 2 ALKALI  
 CC LIGHT CHAIN SUBUNITS (LMC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
 CC CORTEX.  
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN BE FURTHER  
 CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PFM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
 CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE  
 CC ACTIVITY. PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
 CC POSITION (688).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M14628; AAA33227.1; -  
 DR PIR: A26655; A26655.  
 DR PIR: S00250; S00250.  
 DR PDB: 1MAA; 03-DEC-97.  
 DR PDB: 1AMD; 17-AUG-96.  
 DR PDB: 1MMG; 03-DEC-97.  
 DR PDB: 1MMN; 03-DEC-97.  
 DR PDB: 1MND; 17-AUG-96.  
 DR PDB: 1MNE; 17-AUG-96.  
 DR PDB: 1VOM; 23-DEC-96.  
 DR PDB: 1LVK; 28-JAN-98.  
 DR DictyDb: DD01008; mhca.  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRODOM: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IO; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS0096; IQ; 1.  
 KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.  
 FT DOMAIN 1 761  
 FT DOMAIN 762 791  
 FT DOMAIN 817 2116  
 FT NP\_BIND 179 186  
 FT DOMAIN 638 660  
 FT DOMAIN 738 752  
 FT MOD\_RES 130 130  
 FT MOD\_RES 678 678  
 FT MOD\_RES 1823 1823  
 PHOSPHORYLATION (SH-1).  
 PHOSPHORYLATION (BY MHCK).

FT MOD\_RES 1833 1833 PHOSPHORYLATION (BY MHCK).  
 FT MOD\_RES 2029 2029 PHOSPHORYLATION (BY MHCK).  
 SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;  
 Query Match 7.0%; Score 193.5; DB 1; Length 2116;  
 Best Local Similarity 20.8%; Pred. No. 0.37;  
 Matches 140; Conservative 92; Mismatches 257; Indels 185; Gaps 24;  
 QY 9 DGNTSANSADSVGK-----PNLTETSKKTTDSNAVLAVKEVEALLSIDETAAKIG 62  
 DB 964 DGSDPTISRLRKIDELQVEEELTESFSESKRGVL--EKTRVRLQSELDLTVR--- 1018  
 QY 63 KKHQNNLDQFEYHNGSLGAVASIFLIQKQDKGLNEGLKKEIDAKKCSFETNKL 122  
 DB 1019 -----LDSETDKSLLQKKKLEBEILQVOQALAE-----TAAKAOEANKKL 1064  
 QY 123 KEKHTDLQEGVTDADAKKALIKTNGT-----KTGKA----- 154  
 DB 1065 QGEYTELNEKFNSEVTAARSVEKSKKTLQSLQVAVNNELDEKKNRDALKKKALDAML 1124  
 QY 155 EELGKLEPS-----VEYLSKAKKEMLAANSYKELTSFY----- 187  
 DB 1125 EEMKDQLESTGEGKSLYDLKVKQSDMEALRNQISELOSTIAKLEKIKSTLEGEVARLQ 1184  
 QY 188 ----AESPFAMSGNGKGG-----DSASTNPADESAKGPNLTETSKKTTDSNAFVLAVK 236  
 DB 1185 GELAEADQAK-SNEKQKKKVELDLEKSNQALAEETAKQALDKIKKLEB-----LS 1237  
 QY 237 EVELVLVSLDELAKKAIQKIDNNGLAALNNQNSLAGAVASTLTTEKL---SKIKN 293  
 DB 1238 EVQT---OLSEANKNVNSDSTNKHLETFSPNNKLELEAEQKAKQALEKKRLGLESEIKH 1294  
 QY 294 LEE-----LKEITAKKAKCSFEFNKTKSGHADLGKQDATTDHAKA 335  
 DB 1295 VNEDEEKKKOKESNEKRYDLEKEVSELDQITEEYASKAAVTEAKNKESELDITKRO 1354  
 QY 336 ILKHTATTDKAKKEFKDL-----FESVEGL-----KAAQVALTNSYKEIGHR 378  
 DB 1355 YADVSSNRKDSVEQLKTLQAKNEELRMTAEABEQDLRAERSKKKAEFDLEAVKNLEEE 1414  
 QY 379 -----NNSGGGSASINPD-ESAKGN-----LTVISKKTIDS 409  
 DB 1415 TAKKVAEKAKKAKETRYRSKSELDDAKNVSSQYQIKRLNELSELRLVEADERC 1474  
 QY 410 NAFLLAVKEVEALLSI-DELSKAIQKKIKKNDGFLDNEANNEISLIAGAVEISKLITQKL 468  
 DB 1475 NSAIKAKKTAEALSLESLKDEIDANNNAKAKAE-----RKSKLEVRVALESLIEDKS 1527  
 QY 469 SVLNSEELKRIKIRKAKDCSOKFTTKLKDSHAELGIQSVQD-DNAKKAALKTHGTDRGAK 527  
 DB 1528 GTVAVVEFIRKKDAIDDLRLARL-----DRETESIKRDEDEKKNRKQFADLEAKVEEAQR 1582  
 QY 528 E---LEELFKSLES 538  
 DB 1583 EYVTIDRLKRLLES 1596  
 RESULT 11  
 REST\_CHICK STANDARD; PRT; 1433 AA.  
 ID REST\_CHICK  
 AC 042184; 042228; 057563; 057564;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (Clp-170).  
 GN RSN.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]

```

RX SEQUENCE FROM N.A. . Pubmed-9469933:
RA Gripatic L., Volosky J.M., Keller T.C. III;
RT "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RL Gene 206:195-208(1998).
[2]
RN RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC TISSUE=Pectoralis muscle;
RA Gripatic L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms
expressed predominantly in muscle.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
THAT LINKS ENDOCYTIC VESICLES TO MICROBUBLES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CYTOSKELETON (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb.slb.ch).
CC
CC EMBL; AF014012; AACG0344.1; -
DR EMBL; AF020764; AACG0345.1; -
DR EMBL; AF045650; AACG0347.1; -
DR EMBL; AF045651; AACG0348.1; -
DR InterPro; IPR000938; CAP-GLY.
DR InterPro; IPR001878; znf.CCHC.
DR Pfam; PF03302; CAP_GLY_2.
DR SMART; SM00343; znf.C2HC; 1.
DR ProSite; PS00845; CAP_GLY_1; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT DOMAIN 79 121 CAP-GLY 1.
FT FT 144 207 SER-RICH.
FT FT 144 207 CAP-GLY 2.
FT FT 305 332 SER-RICH.
FT FT 351 353 COILED COIL (POTENTIAL).
FT FT 1414 1427 CCHC-BOX.
FT FT 458 492 MISSING (IN SHORT ISOFORM).
FT FT 458 492 TOTALEHRIELOSILEFETKADLORELEDR ->
VARSPLIC 458 492 RKRQISEDPEN (IN ISOFORM CLIP-170(11)).
FT FT 458 492 S -> GGSSSKYS (IN ISOFORM CLIP-170(11)).
FT FT 458 492 T -> RKRQISEDPENT (IN ISOFORM CLIP-
VARSPLIC 458 492 170(11+35)).
FT FT 803 803 K -> R (IN REF. 2; AACG03547).
FT FT 803 803 E -> V (IN REF. 2; AACG03548).
FT FT 440 440 CONFLICT COMPACT
SEQUENCE 1433 AA; 161026 MW; 5631CEB683498E23 CRC64;
-----
Query Match 7.0%; Score 193; DB 1; Length 1433;
Best Local Similarity 22.6%; Pred. No. 0.25;
Matches 137; Conservative 98; Mismatches 280; Indels 92; Gaps 23;
Oy 19 ESVKPNLTETSKITDTSNAVLLAVKEVALLSIDEIAAKIGKKIHONNGLDTREY--- 75
Db 708 ESVEQHDLIVEMEDTLNLKLOEAIEIKVELDVLAQKCNE-QTKLGISLTUOOIRASEEKLDL 766
Oy 76 -----NING-----SLLAGVAISTLKOKLDDGKNE--GLKER-IDAAKCSFT 117
Db 767 AALOKANSEGGKLEIQLESDLOAABEKQIOLETETEVSNLTRELOCKEKKLDLEKNLSA- 825
Oy 118 FTRKLK----EKHTDGRKGVTDA-----DAKEAILKTGTGTGKA---EEEGKLFEE 162
Db 826 -VNQKDSLEKELQLTKRKFTSAVDGAENAORAMDETINKLNQKEOFRLMSSELFOIJS 884
Oy 163 SVEVLAKAEMLANSVKELTSPVAASPPVANGSGNGKGDSTAPNPADS-AKCPNITE 220

```

Db	885	NLMYMERKLEREREREGQLEAVKLVLENDIAELIMKSSGSSQAQLMKMMDLELRKERQEQ	944
Qy	221	ISKRTIDSNAPFLAVK-----EETLVLSIDELAKKAIGOKIDNNGLAALN	267
Db	945	IQLEETKANERKAVALQKNNEQOTQAQKAQSQOETLKTQHEEL-KKMQDOLTTQMKKOMETSQ	10030
Qy	268	NQNSSLIAG-AVAISTLTPEKLSKLNLEELTELAKAKKCEEFETNKLKSGHDLGQD	326
Db	1004	NQYKDLOAKYKEETSEMITKTHDDIGFQONLLDEALKAQAOKKNDELETOAEELKQO	1062
Qy	327	ATDDHAKALIKLTHATDQKAEFKDLPFESVEGLKAAOVALTNSVKEELGHNNSSGCSA	386
Db	1063	--AEQAKA-----DKRAEE--VLOTMEVYTKERAKITQEKITETPLASLENSRQNE	1108
Qy	387	STNDESAKGPNTLVISKRTIDSNAPLV-LAYKEVALLSSIDELSKAIGKIKINDGTLDN	445
Db	1109	KLQNELDMLKONNLKNEELETSTKELLLENKKKEPELAKKPAALKAQAQSQLAAQOE	1168
Qy	446	E-AKRNSLILAGAVEIS-KLTQKLSVLNSE-----ELKKIKKAEADCSQKFTYK	493
Db	1169	ENVKLAEELGRSDREVYSHQKLEERESVYLNQNLLEKKKRESLKKEIDEERSLQK--S	1225
Qy	494	LKDSHAEIGIQSVQDDAKKALIKLTHGTQDQKKELEELFKSLEISIAKQAQALNTSYKE	553
Db	1226	ISDTSLALTQDDELEKLRNETIYVLNG-EWASAKTILQSVYKTLSS---DKLLEEKYKN	1280
Qy	554	LTPNPVA 560	
Db	1281	LEQKLKA 1287	
RESULT	12		
MYSB_CAEEL	ID	MYSB_CAEEL	STANDARD; PRT; 1966 AA.
AC	P02565;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	MYOSIN HEAVY CHAIN B (MHC B).		
GN	UNC-54 OR MYO-4.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;		
OC	Rhabditidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=83273600; PubMed=6576334;		
RA	Karn J., Brenner S., Barnett L.;		
RT	"Protein structural domains in the Caenorhabditis elegans unc-54		
RT	myosin heavy chain gene are not separated by introns.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).		
RN	[2]		
RP	SEQUENCE OF 850-1966 FROM N.A.		
RX	MEDLINE=82272395; PubMed=7202124;		
RA	McLachlan A.D., Karn J.;		
RT	"Periodic charge distributions in the myosin rod amino acid sequence		
RT	match cross-bridge spacings in muscle.";		
RL	Nature 299:226-231(1982).		
RN	[3]		
RP	SEQUENCE OF 1876-1966 FROM N.A.		
RX	MEDLINE=83232892; PubMed=6571695;		
RA	Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,		
RA	Waterston R.H.;		
RT	"The genes sup-7 X and sup-5 III of C. elegans suppress amber		
RT	nonsense mutations via altered transfer RNA.";		
RL	Cell 33:575-583(1983).		
CC	-1- FUNCTION: MUSCLE CONTRACTION.		
CC	-1- SUBUNIT: MUSCLE MYOSIN.		
CC	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)		
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS OF THE MYOFIBRILS.		
CC	-1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.		
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING		
CC	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.		

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC - P1M: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMN). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC - MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.  
 CC ELGANS.  
 CC - MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY  
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC - SIMILARITY: CONTRAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: J01050; AAA28124.1; -  
 CC DR EMBL: V01494; CAA24738.1; -  
 CC DR PIR: A02992; MMKW.  
 CC DR HSSP: P08799; IMND.  
 CC DR InterPro: IPR002928; Myosin\_tail.  
 CC DR InterPro: IPR001609; myosin\_head.  
 CC DR Pfam: PF00063; myosin\_head.  
 CC DR Pfam: PF01576; Myosin\_tail.1.  
 CC DR PRINTS: PRO0193; MYOSINHEAVY.  
 CC DR PRODOM: PD000355; myosin\_head.1.  
 CC DR SMART: SM00242; MYSC: 1.  
 CC KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 CC ATP-binding; Methylation; Alkylation; Multigene family.  
 CC FT DOMAIN 1 850  
 CC FT MYOSIN HEAD-LIKE.  
 CC FT COILED COIL (POTENTIAL).  
 CC FT DOMAIN 851 1164  
 CC FT ALPHA-HELICAL TAILPIECE (S2).  
 CC FT DOMAIN 1165 1176  
 CC FT HINGE.  
 CC FT NP-BIND 1177 1966  
 CC FT LIGHT MEROMYOSIN (LMN).  
 CC FT RC-BIND 184  
 CC FT ATP (BY SIMILARITY).  
 CC FT DOMAIN 655 687  
 CC FT ACTIN-BINDING.  
 CC FT DOMAIN 769 783  
 CC FT METHYLATION (TR1-) (POTENTIAL).  
 CC FT MOD\_RES 128 128  
 CC FT MOD\_RES 705 705  
 CC FT MOD\_RES 715 715  
 CC FT MOD\_RES 705 705  
 CC FT ALKYLATION (SH-1).  
 CC FT ALKYLATION (SH-2).  
 CC FT E -> R (IN REF. 2).  
 CC FT I -> L (IN REF. 2).  
 CC FT CONFLICT 1337 1337  
 CC FT CONFLICT 1880 1880  
 CC SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 7.0%; Score 192; DB 1; Length 1966;  
 Best Local Similarity 21.5%; Pred. No. 0.39; Mismatches 213; Indels 146; Gaps 28;  
 Matches 127; Conservative 104;

DB 16 SADESVKPNLITEISKITDSNAVLAVK---EVEALLSIDET--AAKAIKKIHQ--NN 69  
 DB 1358 SLREEIEGKN--EILQLSKANNDIQWKAREFGEELAA--DELEPAKRROKKNELDQ 1414  
 DB 70 GLDTEYNHNSL-LACAVYAISTLIKOKLDGLKNEGKEIKDAKKCEFTFNKKEKHD 128  
 DB 1415 ALDAANSKNALESKTRSRVGLDLDQAVVERANGVAASALEKKOKGFDKIIDEMRKTTD 1474  
 DB 129 LGEKGTVDADAKFAILKTNGT---KTKGAEELGKLEESVEVSKAKKEMLANSVLELSP 185  
 DB 1475 LAEE---LDGAQRDLRTSTDLFAKNAOE--ELAEEVEGLRRERKS--LSQEIIDLTDQ 1527  
 DB 186 VVAESPAMSGSGKGDASATNPADESANGPNLITEISKITDSNAFVLAKEVETLVLSI 245  
 DB 1528 L-----GEGGRS-----VHMQKIIIRL 1545  
 DB 246 DELAKKAIGOKIDNNNGLAALNNNGSLAGAVYAISTLITEKLSIKNLEELKTEIAR-A 304  
 DB 1546 -EIEKEELQHALDEAE--AALEAESKYLRAQVEVS-----QIRSEIERKRI 1588

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC - P1M: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMN). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC - MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.  
 CC ELGANS.  
 CC - MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY  
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC - SIMILARITY: CONTRAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: J01050; AAA28124.1; -  
 CC DR EMBL: V01494; CAA24738.1; -  
 CC DR PIR: A02992; MMKW.  
 CC DR HSSP: P08799; IMND.  
 CC DR InterPro: IPR002928; Myosin\_tail.  
 CC DR InterPro: IPR001609; myosin\_head.  
 CC DR Pfam: PF00063; myosin\_head.  
 CC DR Pfam: PF01576; Myosin\_tail.1.  
 CC DR PRINTS: PRO0193; MYOSINHEAVY.  
 CC DR PRODOM: PD000355; myosin\_head.1.  
 CC DR SMART: SM00242; MYSC: 1.  
 CC KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 CC ATP-binding; Methylation; Alkylation; Multigene family.  
 CC FT DOMAIN 1 850  
 CC FT MYOSIN HEAD-LIKE.  
 CC FT COILED COIL (POTENTIAL).  
 CC FT DOMAIN 851 1164  
 CC FT ALPHA-HELICAL TAILPIECE (S2).  
 CC FT DOMAIN 1165 1176  
 CC FT HINGE.  
 CC FT NP-BIND 1177 1966  
 CC FT LIGHT MEROMYOSIN (LMN).  
 CC FT RC-BIND 184  
 CC FT ATP (BY SIMILARITY).  
 CC FT DOMAIN 655 687  
 CC FT ACTIN-BINDING.  
 CC FT DOMAIN 769 783  
 CC FT METHYLATION (TR1-) (POTENTIAL).  
 CC FT MOD\_RES 128 128  
 CC FT MOD\_RES 705 705  
 CC FT MOD\_RES 715 715  
 CC FT MOD\_RES 705 705  
 CC FT ALKYLATION (SH-1).  
 CC FT ALKYLATION (SH-2).  
 CC FT E -> R (IN REF. 2).  
 CC FT I -> L (IN REF. 2).  
 CC FT CONFLICT 1337 1337  
 CC FT CONFLICT 1880 1880  
 CC SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 7.0%; Score 192; DB 1; Length 1966;  
 Best Local Similarity 21.5%; Pred. No. 0.39; Mismatches 213; Indels 146; Gaps 28;  
 Matches 127; Conservative 104;

DB 16 SADESVKPNLITEISKITDSNAVLAVK---EVEALLSIDET--AAKAIKKIHQ--NN 69  
 DB 1358 SLREEIEGKN--EILQLSKANNDIQWKAREFGEELAA--DELEPAKRROKKNELDQ 1414  
 DB 70 GLDTEYNHNSL-LACAVYAISTLIKOKLDGLKNEGKEIKDAKKCEFTFNKKEKHD 128  
 DB 1415 ALDAANSKNALESKTRSRVGLDLDQAVVERANGVAASALEKKOKGFDKIIDEMRKTTD 1474  
 DB 129 LGEKGTVDADAKFAILKTNGT---KTKGAEELGKLEESVEVSKAKKEMLANSVLELSP 185  
 DB 1475 LAEE---LDGAQRDLRTSTDLFAKNAOE--ELAEEVEGLRRERKS--LSQEIIDLTDQ 1527  
 DB 186 VVAESPAMSGSGKGDASATNPADESANGPNLITEISKITDSNAFVLAKEVETLVLSI 245  
 DB 1528 L-----GEGGRS-----VHMQKIIIRL 1545  
 DB 246 DELAKKAIGOKIDNNNGLAALNNNGSLAGAVYAISTLITEKLSIKNLEELKTEIAR-A 304  
 DB 1546 -EIEKEELQHALDEAE--AALEAESKYLRAQVEVS-----QIRSEIERKRI 1588

Query Match 6.9%; Score 189; DB 1; Length 1022;  
 Best Local Similarity 23.5%; Pred. No. 0.25;  
 Matches 158; Conservative 89; Mismatches 204; Indels 220; Gaps 37;

```

OY 5 NSGKGNSTNSANADE--SYKGNLTFEISKRITSDSNVLAWEVEALLSIDEIAA-----58
DB 322 NHGQSVALSQTLAEDLTTHVQGPS-HETNRPILPN-----OLESSIEQHTSQOVP 371
OY 59 -----KAIGKKIHQNGDLTEYNHNGSL-----LAGAVAISTLIKOKLDGL-----KNEG 103
DB 372 PTTTNRKSKQPKISQHOHQPOOAGSSGIPNPVLAANALSTSMODLLNINSYLRKND 431
OY 104 LKEKIDAAKKCEFTTNKIK-----EKHTDL-----GKEGVTDADA-----139
DB 432 INKQSDLIKKAIAIAILNNKSDPAEKQYNIIDLAKNIFSNKDIIADAKVAVVLTLETIQ 491
OY 140 -----KEAIIKTNKTKGAELGKLFESVEVLSKAAKEMLANSVKELTSPVAES 190
DB 492 NDQNTLDIRKSKRIEDTVAITLNSENI-ELKQKQDILEKVVD--IGLSIKDISRVAVD 548
OY 191 PAMGNSGKGDASTNPADESAKGNLT-----EIS-----KKITDSNAFLAVREY 238
DB 549 SIMDT-----VKSININEDKEKIFITVFDQINSYFSNAKOKLDS-----ILKTRA 597
OY 239 ETLVLISIDELAKKAIGOKIDN-----NNGI-----AALNNGSLAGAY 278
DB 598 ETQVLSPEQ--QOLMNQNDNITTEHTKRDTEKVNINLEPLSNALAKTTNIOQVTSNV 655
OY 279 AISTLITELSLK-----KNEEL-----KTEIAKAKKCEFTFNKLKSGHAGDQKD 326
DB 656 LQSPVOIEKSKRLIOVTVTAESALVEBKDTETIYK-----IGKTI 698
OY 327 ATDD-----HAKAAILKTHATTDKGAKEFKDFESVEGLLKAQY--ALTNSVKELGHR 378
DB 699 VTHSDTSLPLHDKVYIMGSA-----KGVESKNLDLRELIAGLVDTIYE-----745
OY 379 NNSGDSASTNPDESAGPNLTV--ISKRTDSNAFLAVKEVEALLSIDELS-RAIGK 435
DB 746 -----AKGNNAVVAHAISSMIANSN--INQSEKALKRSQDVYSEKVLDK 787
OY 436 KIKNDCTLDNE--ANRNSLIAG-----AYEISKLITQKLSVLN-----SEE 475
DB 788 ETQN--LDRELKAQININSKLDHDIYNTQOVANALKNVIT--TVLDNSGQSGVSEE 841
OY 476 LKKKIKE-AKDCSOKFTTKLKDASHAELGIQSYODDNAKKAIIKT-HGTRDKGAKELELF 533
DB 842 ARKVVSSLLNDISKRTIEKINLRLMALS-----QDGN-----LKTFEKKDKATKRVDELV 892
OY 534 KSELESTSKAAQ 544
DB 893 KAFDNKSTEE 903

```

```

RA Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA Cerletti N., Sorg C., Odink K., Tarcsey L., Wiesendanger W.,
RA de Wolf-Peters C., Shipman R.,
RA Redlin: a novel intermediate filament-associated protein highly
RA expressed in the Reed-Sternberg cells of Hodgkin's disease."
RA EMBO J. 11:2103-2113(1992).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-92405160; PubMed-1356075;
RX Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
RA "CLIP-170 links endocytic vesicles to microtubules.";
RL Cell 70:887-900(1992).
CC CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
CC CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CYTOSKELETON.
CC CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
CC OF HODGKIN'S DISEASE.
CC CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X64838; CAA46050.1; -
DR EMBL: M97501; AAA35693.1; -
DR PIR: S22695; S22695.
DR MIM: 179838; -
DR InterPro: IPR000938; CAP-GLY.
DR InterPro: IPR001878; znf.CCHC.
DR Pfam: PF01302; CAP-GLY; 2.
DR SMART: SM00343; znf.C2HC; 1.
DR PROSITE: PS00845; CAP-GLY_1; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT DOMAIN 60 125 CAP-GLY 1.
FT FT 125 204 SER-RICH.
FT FT 204 279 CAP-GLY 2.
FT DOMAIN 304 331 SER-RICH.
FT DOMAIN 350 1342 COILED COIL (POTENTIAL).
FT DOMAIN 1408 1421 CCHC-BOX.
FT VARSPLIC 457 491 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1069 1069 D -> E (IN REF. 2).
SQ :SEQUENCE 1427 AA; 160989 MW; 0A4F166DD94254E8 CRC64;

```

Query Match 6.9%; Score 189; DB 1; Length 1427;  
 Best Local Similarity 21.6%; Pred. No. 0.36;  
 Matches 135; Conservative 100; Mismatches 249; Indels 140; Gaps 26;

```

OY 28 EISKKITDSNAVLAVKEVEALLSIDELAKAIGKKIHQNGID-----TEYNHNGS 80
DB 523 ESNKPGADVMSLSLLOEISSLOEKLE-----VTRDHOREITSLKEHNGAREETHQKE 576
OY 81 LLAGAVAISTLIKOKLDGLKNEGKKEIDAAKKCEFTTNKLEK-HTDQKSGVTDADA 139
DB 577 IKALVYATEKLSKE-----NESIKSKLEHANNKNSDVIALMKSKLETAISHQOAMEEL 630
OY 140 KEAIIKTNKTKGAELGKLFESVEVLSKAAKEMLANSVKELTSPVAESPAWGSN---196
DB 631 KVSFSGIGLGTETAFELKQIETKMRDYOHEIENLQNDQSEBAHAKEMELAKLAKLM 690
OY 197 --SGKGDASTNPADESAGPNLTETSKRTITDSNAFLAVKEVEVTLVLSIDELAK-----250
DB 691 VIREKENSLEAIRSKLDKADQHLVEMEDTLNKQEAIEIVKTEVLAQAKCNQDTKVIDN 750
OY 251 -----KAIGOKI-----DNNNGLAALNNQNGSLAGAVAI STL-----I 284

```

```

Db 751 FTSQKATEKELLDLALRKASSEKSEKUKLQOUEAAEKOIKHLEIEKNASSKASSI 810
QY 285 TEKLS-----KLNLEIKTEIKAKKSCSEPTNKKLSGHADIGKODATDHA--KKAIIKT 339
Db 811 TRELOREKELKTNLOJENLESEVQVETLEKELOILKEKKAESSEAVSORSMOETVNL 870
QY 340 HA-----TTDKAKERKDLFESEVGLLKAQVALTNSVKEIGHRNNS 381
Db 871 HKEEQFNMLSSDLEKRLRENLMDEAK--FREKDEREOLITKAKE--KLENDIAEI---MKM 925
QY 382 GGDAS-----TNPDESAGPNLTIVISKTTIDSNAPLAVKEVALISIDELSKAIGKKI 437
Db 926 SGDNSQSLTKMDELKRLKRDVEELQKLTJKN-----ENASFLQKSTIEDWT-----V 973
QY 438 KNDGTLDNEANRNESLIGAVEISKLITQKLSVLNSEELKKI-----KRAKCSQKF 490
Db 974 KAEQSQOQEAARKHE-----EKKELERKLS-----DLEKMEISHNOCQOLKARYERA 1021
QY 491 TTKLDSHALEIGSQVODNNAKALIKTHGTOKGAK-----ELELEFKSLLESLSKA 542
Db 1022 TSETTKRHEI--LQNLQ-----KTLIDTE-DKLGKARENSGLLOLEBELRKQADK--AKA 1073
QY 543 AQA-----LTNSVKELTNPYVA 560
Db 1074 AQTADDAQIMEQMKREKETTLAS 1097

```

```

RESULT 15
Z1PL_YEAST STANDARD: PRT: 875 AA.
AC P31111:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SYNAPTONEMAL COMPLEX PROTEIN Z1PL.
GN Z1PL OR YDR285W OR D9819.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BR1824-3B;
RX MEDLINE=93161412; PubMed=7916652;
RA Sym M., Engelbrecht J.A., Roeder G.S.;
RA "Z1PL is a synaptonemal complex protein required for meiotic
chromosome synapsis."
RA Cell 72:365-378(1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnson M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Pavello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifken L., Riles L., Taich A., Trevisan E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RL
CC -!- FUNCTION: REQUIRED FOR MEIOTIC CHROMOSOME SYNAPSIS AND CELL CYCLE
PROGRESSION. MAY ACT AS A MOLECULAR ZIPPER TO BRING HOMOLOGOUS
CHROMOSOMES IN CLOSE APPPOSITION. Z1PL MAY ENCODE THE TRANSVERSE
FILAMENTS OF THE SYNAPTONEMAL COMPLEX.
CC -!- SUBCELLULAR LOCATION: SYNAPSED MEIOTIC CHROMOSOMES.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

```

DR EMBL: L06487; AAA35239.1; -;
DR EMBL: U51031; AAB64474.1; -;
DR PIR: A45173; A45173.
DR PIR: S30868; S30868.
DR SGD: S0002693; Z1PL.
KW Nuclear protein; Meiosis; Coiled coil.
FT DOMAIN 177 333 COILED COIL (POTENTIAL).
FT DOMAIN 397 438 COILED COIL (POTENTIAL).
FT DOMAIN 456 752 COILED COIL (POTENTIAL).
FT CONFILCT 55 55 T -> A (IN REF. 1).
SQ SEQUENCE 875 AA; 100035 MW; 674F12625CD9DDFD CRC64;

```

Query Match 6.8%; Score 187; DB 1; Length 875;

Best Local Similarity 20.18; Pred. No. 0.25; Matches 115; Conservative 117; Mismatches 197; Indels 142; Gaps 25;

```

QY 20 SVKGPNTL--EISKRTDSNAVLAV--KEYEALLSIDETIAAKAGKRIHQNGGLDTEYN 76
Db 274 STEMLNTINDLGKRRKEADAELEKKGKEIYELKRELDGCSGLSEKTI--KNSSLIQEMG 332
QY 77 HNSLLAGATAISTLIKOKIDLGKNEGKLEKI-----DAKKCSEPTNKKLKEKHTDI 129
Db 333 KNR-----EEMKRSIENFESEDKAHLLQFNKFEERVHDLFEKKL--QKHFDV 378
QY 130 GKEGVTDADAKKALIKNGTKTGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVAE 189
Db 379 AKDTLNGGLNTYVELSSNTETMLKQOYEDIKENLEOKMSSKDEAKTITNELS--VTOK 436
QY 190 SPANGSGSGKGGDSASTNPDESAGKPNLTETISKTTDSNAFVLAVKEVETLVSIDELA 249
Db 437 GLIMG-----VOEELITSG-----NIOTALVS----- 459
QY 250 KKAIGKIDNNGGLALNNNGSLAGAVASITLITKSLKNT--ELKTELAKKKCS 308
Db 460 -----EKNNTROELDDA-----SQTAKNYSLENLKVAKAEIVOSNE--Y 499
QY 309 EEFNTKLSGHADLGKQDATDHAKAAILKTHATTDKAKE--FKDLFESVEGLKKAQV 366
Db 500 EERIKHLESERSTLSSQ-----KNQIISLGRKEAQYEDLVAKLEA--KNIEI 545
QY 367 A-LTNSVKELGHNNGSGDSASTNPDESAGKPNLTIVISK-----KITDSNAFLAVKEVE 420
Db 546 SQISGKEQSLTEKENNLSNELKRVQDQLEKLNLTITTSYENKISSONEI-----YK 599
QY 421 ALLSIDELSKAIGKKI-----KNDGTLDNEA--NRNSLIGAVEISKLITQKLSVLN 472
Db 600 ALVSENDTLKRIQQLVEIKENQDKHTTKLEAFQKNNEQLOKLVAVEVQLAHLEL----- 655
QY 473 SEELKKIKKAEKDCSQKFTTKLKDSHAELGI--QSV-----QDNMAKALIKTHGTRD 523
Db 656 --ELEBQNHKLKNCLEKKEKTEGVEESLSDVTKLQOYIVLKSQDITAEKLELQON----- 709
QY 524 KGAKLEELFKSLLESLSKAAQALTNVSKEL 554
Db 710 --LESLEVTKNLQOKVQVQKRELEOKIKEL 738

```

Search completed: March 18, 2002, 10:12:03  
Job time: 981 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:55 ; Search time 124.19 seconds

(without alignments)  
659.574 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750  
Sequence: 1 MACNNSGKNGNTSANSADAES.....KAAQALITNSVKELTNPVVA 560

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SPREMBL.17:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phage:\*
- 11: sp\_plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp\_vertebrate:\*
- 15: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	926	33.7	192	2	Q9S3P3
2	921	33.5	192	2	Q9R7B1
3	910	33.1	191	2	P70818
4	908	33.0	193	2	Q9R853
5	893	32.5	200	2	Q9R089
6	886	32.2	194	2	Q9S3P4
7	885	32.2	182	2	Q9R7B2
8	884	32.1	203	2	P96507
9	881	32.0	212	2	Q07336
10	880	32.0	207	2	Q9KIM4
11	878	31.9	207	2	Q45187
12	878	31.9	207	2	Q9REF8
13	878	31.9	207	2	Q9REF7
14	877	31.9	212	2	Q44669
15	876	31.9	203	2	Q50623
16	874	31.8	200	2	Q08139
17	853	31.0	177	2	Q44999
18	846	30.8	179	2	Q44987
19	844	30.7	179	2	Q44988

20	843.5	30.7	178	2	Q9REH7	Q9REH7 borrelia bu
21	834.5	30.3	180	2	Q44991	Q44991 borrelia af
22	832	30.3	179	2	Q44986	Q44986 borrelia af
23	829	30.1	175	2	Q9R7B4	Q9R7B4 borrelia bu
24	805	29.3	189	2	P94224	P94224 borrelia bu
25	802	29.2	163	2	Q9R7A9	Q9R7A9 borrelia bu
26	798	29.0	207	2	Q45177	Q45177 borrelia ga
27	794	28.9	200	2	Q44996	Q44996 borrelia ga
28	786	28.6	207	2	Q49581	Q49581 borrelia ga
29	773	28.1	159	2	Q9R7B0	Q9R7B0 borrelia bu
30	759.5	27.6	199	2	Q44998	Q44998 borrelia ga
31	757.5	27.5	211	2	Q44720	Q44720 borrelia bu
32	756.5	27.5	193	2	P94237	P94237 borrelia bu
33	755.5	27.5	199	2	Q08233	Q08233 borrelia ta
34	745.5	27.1	204	2	P96509	P96509 borrelia af
35	705.5	25.7	206	2	P96512	P96512 borrelia af
36	700	25.5	210	2	Q44719	Q44719 borrelia bu
37	698	25.4	192	2	Q9S3P2	Q9S3P2 borrelia bu
38	696.5	25.3	193	2	P94234	P94234 borrelia bu
39	691	25.1	193	2	Q31115	Q31115 borrelia bu
40	687.5	25.0	204	2	Q44989	Q44989 borrelia af
41	683.5	24.9	211	2	Q49576	Q49576 borrelia af
42	678.5	24.7	194	2	Q31122	Q31122 borrelia bu
43	678.5	24.7	211	2	Q44977	Q44977 borrelia bu
44	677	24.6	212	2	Q9KIM5	Q9KIM5 borrelia af
45	676.5	24.6	209	2	Q44671	Q44671 borrelia af

ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
AC	Q9S3P3	01-MAY-2000 (TREMELREL. 13, Created)			
DT	01-MAY-2000 (TREMELREL. 13, Last sequence update)				
DT	01-JUN-2001 (TREMELREL. 17, Last annotation update)				
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IP2;				
RX	MEDLINE=96296448; PubMed=8709845;				
RA	Lively I., Gibbs C.P., Schuster R., Dornier F.;				
RT	"Evidence for lateral transfer and recombination in OspC variation in				
RT	Lyme disease Borrelia."				
RL	Mol. Microbiol. 18:257-269(1995).				
DR	EMBL, L42887; AAB36995.1; "				
DR	InterPro: IPR001800; Lipoprotein_6.				
DR	Pfam: PF01441; Lipoprotein_6; 1.				
FT	PRODOM: PD001149; Lipoprotein_6; 1.				
FT	NON_TER 1 192				
FT	NON_TER 1 192				
SQ	SEQUENCE 192 AA; 20287 MW; 11846F7AC84C7E3D CRC64;				
Query Match 33.7%; Score 926; DB 2; Length 192;					
Best Local Similarity 99.5%; Pred. No. 3.9e-35;					
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
OY	3	CNNNSGKGNTSANSADAESVKGPNLTETSKRTIPDSNVLLAVKVEALLSIDIAKAIG 62			
DB	1	CNNNSGKGNTSANSADAESVKGPNLTETSKRTIPDSNVLLAVKVEALLSIDIAKAIG 60			
OY	63	KTIHNNGLDTEVYHNHNSLAGAVALSTLIKOKLDIKNEGLKEKIDAAKCKSEFTTNKL 122			
DB	61	KTIHNNGLDTEVYHNHNSLAGAVALSTLIKOKLDIKNEGLKEKIDAAKCKSEFTTNKL 120			
OY	123	KEKHTDLGREGYVDADAKKELIKTNGTKTGAEELGKIFESVSVLSKAKKEMLVSKVEL 182			

```

Db 121 KEKHTDLGKRGVTDADAKKAIKTKNGTKTGAEELGKLFESVEVLSKAKEMLANSVKEL 180
QY 183 TSPVAESP 191
    |||||
Db 181 TSPVAESP 189

RESULT 2
O9R7B1
ID 09R7B1 PRELIMINARY: PRT: 192 AA.
AC 09R7B1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

[1]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RN U.S.A.'?",
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RM [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RM EMBL: U91798; AAB81895.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AAAF64 CRC64;

Query Match 33.5%; Score 921; DB 2; Length 192;
Best Local Similarity 98.4%; Pred. No. 6.6e-35;
Matches 187; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKA 60
    :|||||
    3 ISCNNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKA 62
OY 61 IGKRIHONGIDTEYNNHNSILAGAVAI STLIRKOKLDGKNEGKKEKIDAAKCSFTFN 120
    |||||
Db 63 IGKRIHONGIDTEYNNHNSILAGAVAI STLIRKOKLDGKNEGKKEKIDAAKCSFTFN 122
OY 121 KLAKEKHTDLGKRGVTDADAKKAIKTKNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
    |||||
Db 123 KLAKEKHTDLGKRGVTDADAKKAIKTKNGTKTGAEELGKLFESVEVLSKAKEMLANSVK 182
OY 181 ELPSPVAES 190
    |||||
Db 183 ELPSPVAES 192

RESULT 3
P70818
ID P70818 PRELIMINARY: PRT: 191 AA.
AC P70818
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DE OUTER SURFACE PROTEIN (FRAGMENT).
OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2-1498 CAD;
RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: L81131; AAB06569.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 191
SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C62B4DC3C0 CRC64;

Query Match 33.1%; Score 910; DB 2; Length 191;
Best Local Similarity 98.9%; Pred. No. 2.1e-34;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 NNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKAIGK 63
    |||||
Db 1 NNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKAIGK 60
OY 64 KIHONGIDTEYNNHNSILAGAVAI STLIRKOKLDGKNEGKKEKIDAAKCSFTFNK 123
    |||||
Db 61 KIHONGIDTEYNNHNSILAGAVAI STLIRKOKLDGKNEGKKEKIDAAKCSFTFNK 120
OY 124 EKHTDLGKRGVTDADAKKAIKTKNGTKTGAEELGKLFESVEVLSKAKEMLANSVKELT 183
    |||||
Db 121 EKHTDLGKRGVTDADAKKAIKTKNGTKTGAEELGKLFESVEVLSKAKEMLANSVKELT 180
OY 184 SPVAESP 191
    |||||
Db 181 SPVAESP 188

RESULT 4
O9R53
ID O9R53 PRELIMINARY: PRT: 193 AA.
AC O9R53
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OC1;
RX MEDLINE=99091544; PubMed=9872945;
RA Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of ospC in a local population of Borrelia
RN burgdorferi sensu stricto.",
RL Genetics 151:15-30(1999).
DR EMBL: AF029860; AAB86543.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20502 MW; 5EFD85AF8986D1E CRC64;

Query Match 33.0%; Score 908; DB 2; Length 193;
Best Local Similarity 98.4%; Pred. No. 2.6e-34;
Matches 184; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKA 60
    :|||||
    7 ISCNNSGKDGNTSANSADSEYKGNLTETISKRTDSNAVLAVKEVALLSSIDEIAKA 66
OY 61 IGKRIHONGIDTEYNNHNSILAGAVAI STLIRKOKLDGKNEGKKEKIDAAKCSFTFN 120

```

```

DB 67 ICKKHONNGDTEENNHNHSLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 126
OY 121 KIKERHTDLGKGVTDADAKKAI LKTNKTKTGAEELGKLFESEVLSAAKEMLANSVK 180
DB 127 KIKERHTDLGKGVTDADAKKAI LKTNKTKTGAEELGKLFESEVLSAAKEMLANSVK 186
OY 181 ELTSPV 187
DB 187 ELTSPV 193

RESULT 5
O9ROR9 PRELIMINARY: PRT: 200 AA.
AC O9ROR9:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).
Plasmid cp26.
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2E7;
RX MEDLINE=20002545; PubMed=10531219;
RA Homeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT "Population dynamics of a naturally occurring heterogeneous mixture of
  Borrelia burgdorferi clones."
RL Infect. Immun. 67:5709-5716(1999).
DR EMBL: AF074464; AAD3911.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KM Plasmid.
FT NON_TER 200
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BBEB6B CRC64;

Query Match 32.5%; Score 893; DB 2; Length 200;
Best Local Similarity 98.4%; Pred. No. 1.3e-33;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKGGNTSANSADSVKGPNLTEISKRTTDSNAVLAVKEVALLSIDETAAKA 60
DB 17 ISCNNSGKGGNTSANSADSVKGPNLTEISKRTTDSNAVLAVKEVALLSIDETAAKA 76
OY 61 ICKKHONNGDTEENNHNHSLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 120
DB 77 ICKKHONNGDTEENNHNHSLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 136
OY 121 KIKERHTDLGKGVTDADAKKAI LKTNKTKTGAEELGKLFESEVLSAAKEMLANSVK 180
DB 137 KIKERHTDLGKGVTDADAKKAI LKTNKTKTGAEELGKLFESEVLSAAKEMLANSVK 196
OY 181 ELTSPV 184
DB 197 ELTSPV 200

RESULT 6
O9S3P4 PRELIMINARY: PRT: 194 AA.
AC O9S3P4:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).

```

```

OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JSB;
RX MEDLINE=96296448; PubMed=8709845;
RA Livey I., Gibbs C.P., Schuster R., Dörner F.;
RT "Evidence for lateral transfer and recombination in OspC variation in
  Lyme disease Borrelia."
RL Mol. Microbiol. 18:257-269(1995).
DR EMBL: L42883; AAB36991.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 194
SQ SEQUENCE 194 AA; 20446 MW; CEEDC9FA5DF0D68F CRC64;

Query Match 32.2%; Score 886; DB 2; Length 194;
Best Local Similarity 90.7%; Pred. No. 2.5e-33;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

OY 195 SNSGKGSASSTNPADESA KGPVLTEISKRTTDSNAFLAVKEVTLVLSIDELAKKAIG 254
DB 2 NNSGKGSASSTNPADESA KGPVLTEISKRTTDSNAFLAVKEVTLVLSIDELAKKAIG 61
OY 255 ORIDNNNGSLAALNONGSLAGAVAI STLTEKSLKLNLEELKTEIAKAKCCSEFTNK 314
DB 62 ORIDNNNGSLAALNONGSLAGAVAI STLTEKSLKLNLEELKTEIAKAKCCSEFTNK 121
OY 315 LKSGHADLKODATDDHAKAAILKTHATDKGAKKEDLFESEVGLKKAQVALTNSVKE 374
DB 122 LKSGHADLKODATDDHAKAAILKTHATDKGAKKEDLFESEVGLKKAQVALTNSVKE 181
OY 375 LGRNNSGGSASTNP--DESAKGP 397
DB 182 L-----TSPVAESPKKP 194

RESULT 7
O9R7B2 PRELIMINARY: PRT: 182 AA.
AC O9R7B2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?"
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB81894.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 182
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

```

Query Match 32.2%; Score 885; DB 2; Length 182;  
Best Local Similarity 99.5%; Pred. No. 2.6e-33;  
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 7 GKQNTSANSADSVKPNLITEISKITDSNAVLAVKEVEALLSIDELAIAIKKIH 66
DB 1 GKQNTSANSADSVKPNLITEISKITDSNAVLAVKEVEALLSIDELAIAIKKIH 60
QY 67 QNNGLDTEYHNHNSLAGAVAIISTLLKOKLDGKNGEIKEDIAAKCSEFTNKLKEKH 126
DB 61 QNNGLDTEYHNHNSLAGAVAIISTLLKOKLDGKNGEIKEDIAAKCSEFTNKLKEKH 120
QY 127 TDGKRGVYDADAKKAIKLTNGTKGAEELKLFESVEVLSKAKEMLANSVKELTSPV 186
DB 121 TDGKRGVYDADAKKAIKLTNGTKGAEELKLFESVEVLSKAKEMLANSVKELTSPV 180
QY 187 VA 188
DB 181 VA 182
```

## RESULT 8

```
ID P96507 PRELIMINARY; PRT: 203 AA.
AC P96507:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSpC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BXF;
RA Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB000985; BAA24125.1; -
DR EMBL: AB000985; BAA24125.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 203
SQ SEQUENCE 203 AA; 21549 MW; 94A865FBBACEF1C4 CRC64;
```

Query Match 32.1%; Score 884; DB 2; Length 203;  
Best Local Similarity 99.4%; Pred. No. 3.3e-33;  
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 195 SNSGKGDSASTNPADSAGPNLITEISKITDSNAFLAVKEVELVLSIDELAKKAIG 254
DB 14 SNSGKGDSASTNPADSAGPNLITEISKITDSNAFLAVKEVELVLSIDELAKKAIG 73
QY 255 OKIDNNNGSLAALNNONGSLAGAVAIISTLLTERLSKLNLEELKTELAKRKKCSEFTNK 314
DB 74 OKIDNNNGSLAALNNONGSLAGAVAIISTLLTERLSKLNLEELKTELAKRKKCSEFTNK 133
QY 315 LKSGHADLCKQDATDHAKAAILKTHATDCKAKKEFKDLFESVEGLLKAQVALTNSVKE 374
DB 134 LKSGHADLCKQDATDHAKAAILKTHATDCKAKKEFKDLFESVEGLLKAQVALTNSVKE 193
QY 375 L 375
DB 194 L 194
```

## RESULT 9

```
ID 007336 PRELIMINARY; PRT: 207 AA.
AC 007336:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C.
GN OSpC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB1;
RA MEDLINE-93268136; PubMed-8098841;
RA Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA Will G., Wilske B.;
RT "Genetic heterogeneity of the genes coding for the outer surface protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL Med. Microbiol. Immunol. 182:37-50(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PB1;
RA MEDLINE-93293332; PubMed-8478108;
RA Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA Schwab E., Wanner G.;
RT "Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of Borrelia burgdorferi.";
RL Infect. Immun. 61:2182-2191(1993).
DR EMBL: X69595; CAA49305.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
SQ SEQUENCE 207 AA; 22320 MW; 4A525ABDAE63B5C8 CRC64;
```

Query Match 32.0%; Score 881; DB 2; Length 207;  
Best Local Similarity 99.5%; Pred. No. 4.6e-33;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 379 SNSGGSASASTNPDESAGPNLITVYSKKITDSNAFLAVKEVEALLSIDELSAIGKTKR 438
DB 20 SNSGGSASASTNPDESAGPNLITVYSKKITDSNAFLAVKEVEALLSIDELSAIGKTKR 79
QY 439 NDCITLNEANRNESLJAGAVEISKLTITOKLSVNSELEKRIKEAKDCSQKFTTKLDKSH 498
DB 80 NDCITLNEANRNESLJAGAVEISKLTITOKLSVNSELEKRIKEAKDCSQKFTTKLDKSH 139
QY 499 AELGISVODDNKKAAILKTHGKDGAKELBELFRSLSELSKAAQALTNVSKELTNPV 558
DB 140 AELGISVODDNKKAAILKTHGKDGAKELBELFRSLSELSKAAQALTNVSKELTNPV 199
QY 559 VA 560
DB 200 VA 201
```

## RESULT 10

```
ID 09K1M4 PRELIMINARY; PRT: 212 AA.
AC 09K1M4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C.
GN OSpC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-PCAU:  
RA Dykhuizen D.E., Gutman D.S., Luft B.J.:  
RT "Antigenic variation and integrative recombination in the ospC gene of  
RT Borrelia burgdorferi.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF230185; AAF75626.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
SQ SEQUENCE 212 AA; 22495 MW; C51DAF803BF2E7D4 CRC64;

Query Match 32.0%; Score 880; DB 2; Length 212;  
Best Local Similarity 90.2%; Pred. No. 5.2e-33;  
Matches 185; Conservative 2; Mismatches 4; Indels 14; Gaps 2;

QY 195 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELAKKAIG 254  
DB 20 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELAKKAIG 79

QY 255 OKIDNNNGSLAALNONGSLAGAYASTLITETKSLKMLELKTETIAKAKCSEFTTK 314  
DB 80 OKIDNNNGSLAALNONGSLAGAYASTLITETKSLKMLELKTETIAKAKCSEFTTK 139

QY 315 LKSGHADGKODATDPAKATILKTHATDKGAKKEFKDLFESVEGLKKAQVALTNSVKE 374  
DB 140 LKSGHADGKODATDPAKATILKTHATDKGAKKEFKDLFESVEGLKKAQVALTNSVKE 199

QY 375 LGRNNSGDSASTNP--DESAKGP 397  
DB 200 L-----TSPVVAESPKEP 212

RESULT 11  
Q45187 PRELIMINARY; PRT; 207 AA.

AC Q45187;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN C PRECURSOR.  
GN OSPC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_Taxid:29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PTROB;  
RX MEDLINE-95213332; PubMed-7699024;  
RA Jilka B., Jauris-Helpe S., Lobentanz R., Pradel I.,  
RA Preac-Mursic V., Roseler D., Soutschek E., Johnson R.C.;  
RA Phenotypic analysis of outer surface protein C (OSPC) of Borrelia  
RT burgdorferi sensu lato by monoclonal antibodies: relationship to  
RT genospecies and OspA serotype.";  
RL J. Clin. Microbiol. 33:103-109(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PTROB;  
RX MEDLINE-95395018; PubMed-7655660;  
RA Jauris-Helpe S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,  
RA Soutschek E., Will G., Wilske B.;  
RT "Molecular analysis of genes encoding outer surface protein C (OSPC)  
RT of Borrelia burgdorferi sensu lato: relationship to ospA genotype and  
RT evidence of lateral gene exchange of ospC.";  
RL J. Clin. Microbiol. 33:1860-1866(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PCSC, PBAETI, PFIL, PMUE, PSB;  
RA Marconi R.T., Hohenberger S., Jauris-Helpe S., Schulte-Spechtel U.,  
RA Lavole C.P., Roessler D., Wilske B.;  
RT "Genetic analysis of B.garinii OspA-serotype 4 strains associated with  
RT neuroborreliosis: evidence for extensive genetic homogeneity.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: X83554; CAA58544.1; -  
DR EMBL: AJ236908; CAB46238.1; -  
DR EMBL: AJ132793; CAB46231.1; -  
DR EMBL: AJ132796; CAB46234.1; -  
DR EMBL: AJ132797; CAB46235.1; -  
DR EMBL: AJ132798; CAB46236.1; -  
DR EMBL: AJ236907; CAB46237.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
KW Signal.  
FT CHAIN 1 18 POTENTIAL.  
SQ SEQUENCE 207 AA; 22321 MW; 20889AA76E63AA9D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;  
Best Local Similarity 98.9%; Pred. No. 6.2e-33;  
Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELSKAIGKIK 438  
DB 20 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELSKAIGKIK 79

QY 439 NOSTLNEANRNSLJAGAYEISKLITOKLSVNSEELKKIKKADCSOKFTTKLKD 498  
DB 80 NOSTLNEANRNSLJAGAYEISKLITOKLSVNSEELKKIKKADCSOKFTTKLKD 139

QY 499 AELGISOVDNNAKKAIIKTHGTRDKGAKLEELFSLSLSKRAQAALNSVKELTNPV 558  
DB 140 AELGISOVDNNAKKAIIKTHGTRDKGAKLEELFSLSLSKRAQAALNSVKELTNPV 199

QY 559 VA 560  
DB 200 VA 201

RESULT 12  
Q9REF8 PRELIMINARY; PRT; 207 AA.

AC Q9REF8;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OUTER SURFACE PROTEIN (FRAGMENT).  
GN OSPC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_Taxid:29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PWA;  
RX MEDLINE-20033636; PubMed-10565915;  
RA Marconi R.T., Hohenberger S., Jauris-Helpe S., Schulte-Spechtel U.,  
RA Lavole C.P., Roessler D., Wilske B.;  
RT "Genetic analysis of Borrelia garinii OspA serotype 4 strains  
RT associated with neuroborreliosis: evidence for extensive genetic  
RT homogeneity.";  
RL J. Clin. Microbiol. 37:3965-3970(1999).  
DR EMBL: AJ132794; CAB46232.1; -  
DR InterPro: IPR001800; Lipoprotein\_6.  
DR Pfam: PF01441; Lipoprotein\_6; 1.  
DR ProDom: PD001149; Lipoprotein\_6; 1.  
FT NON-TER 207 207  
SQ SEQUENCE 207 AA; 22366 MW; 30CFEAA6D913A49D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;  
Best Local Similarity 98.9%; Pred. No. 6.2e-33;  
Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELSKAIGKIK 438  
DB 20 NSNGSGDSASTNPDESAGPMLTETISKRITDSNAFLAVKEVETLVLSIDELSKAIGKIK 79

```

DB 20 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIGKIK 79
QY 439 NDGTLNDEANRNESILAGAVEISKLTOKLSVNSEELKKIKKADCKSOFTTKLDSH 498
DB 80 NCGTLNDEANRNESILAGAVEISKLTOKLSVNSEELKKIKKADCKSOFTTKLDSH 139
QY 499 AELGISVODDANKAKILKTHGTDKGAKELEELFKSLESLSKAQAALNTSVKELTNPV 558
DB 140 AELGISVODDANKAKILKTHGTDKGAKELEELFKSLESLSKAQAALNTSVKELTNPV 199
QY 559 VA 560
DB 200 VA 201

RESULT 13
Q9REF7 PRELIMINARY: PRT: 207 AA.
Q9REF7:
01-MAY-2000 (TREMBLrel. 13, Created)
01-JUN-2001 (TREMBLrel. 17, Last sequence update)
OUTER SURFACE PROTEIN.
OSPC.
Borrelia garinii.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=29519;
SEQUENCE FROM N.A.
STRAIN-PHOE;
MEDLINE=20033636; PubMed=10565915;
Marconi R.T., Hoenesberger S., Jauris-Helpe S., Schulte-Spechtel U.,
LaVoie C.P., Roessler D., Wilske B.;
"Genetic analysis of Borrelia garinii Ospa serotype 4 strains
associated with neuroborreliosis evidence for extensive genetic
homogeneity.";
J. Clin. Microbiol. 37:3965-3970(1999).
EMBL: AJ132795; CAB6233.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
SEQUENCE 207 AA; 22249 MW; 20899AA76E0C49D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;
Best Local Similarity 98.9%; Pred. No. 6.2e-33;
Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

379 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIGKIK 438
DB 20 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIGKIK 79
QY 439 NDGTLNDEANRNESILAGAVEISKLTOKLSVNSEELKKIKKADCKSOFTTKLDSH 498
DB 80 NDGTLNDEANRNESILAGAVEISKLTOKLSVNSEELKKIKKADCKSOFTTKLDSH 139
QY 499 AELGISVODDANKAKILKTHGTDKGAKELEELFKSLESLSKAQAALNTSVKELTNPV 558
DB 140 AELGISVODDANKAKILKTHGTDKGAKELEELFKSLESLSKAQAALNTSVKELTNPV 199
QY 559 VA 560
DB 200 VA 201

RESULT 14
Q44669 PRELIMINARY: PRT: 212 AA.
Q44669:
01-NOV-1996 (TREMBLrel. 01, Created)
01-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OUTER SURFACE PROTEIN C.

```

```

GN OSPC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PHO;
RX MEDLINE=95395018; PubMed=7665660;
RA Jauris-Helpe S., Ilegi G., Preac-Mursic V., Roessler D., Schwab E.,
Rautschek E., Will G., Wilske B.;
"Genetic analysis of genes encoding outer surface protein C (OSPC)
of Borrelia burgdorferi sensu lato: relationship to ospa genotype and
evidence of lateral gene exchange of ospC.";
J. Clin. Microbiol. 33:1860-1866(1995).
EMBL: X81521; CA57241.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
SEQUENCE 212 AA; 22514 MW; C6EDC3BFD488DDE CRC64;

Query Match 31.9%; Score 877; DB 2; Length 212;
Best Local Similarity 89.8%; Pred. No. 7.1e-33;
Matches 184; Conservative 4; Mismatches 3; Indels 14; Gaps 2;

QY 195 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIG 254
DB 20 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIG 79
QY 255 QKIDNNGALANNGSLAGAVASTLTERLSKLNKEELTEETAKAKKCEFTTNK 314
DB 80 QKIDNNGALANNGSLAGAVASTLTERLSKLNKEELTEETAKAKKCEFTTNK 139
QY 315 LKSHADLGGODADDEHAKAKILKTHATTDKGAKEFDIESVGLKKAQVALTNSVKE 374
DB 140 LKSHADLGGODADDEHAKAKILKTHATTDKGAKEFDIESVGLKKAQVALTNSVKE 199
QY 375 LGRNNSGDSASTNP--DESAKGP 397
DB 200 L-----TSPIVVAESPKKP 212

RESULT 15
Q50623 PRELIMINARY: PRT: 203 AA.
Q50623:
01-JUN-1998 (TREMBLrel. 06, Created)
01-JUN-1998 (TREMBLrel. 06, Last sequence update)
01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PWUD;
RA Masuzawa T., Sawaki K., Yoshii T.;
Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
EMBL: AB009899; BAA24129.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 203
FT NON_TER 1
SEQUENCE 203 AA; 21505 MW; 66DF78BBAE219 CRC64;

Query Match 31.9%; Score 876; DB 2; Length 203;
Best Local Similarity 98.9%; Pred. No. 7.5e-33;
Matches 179; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 195 NNSGDSASTNPDESAGKPNLTVISKTTDSNAFLAVKEVALSSIDELSKAIG 254

```

```

Db      14  NNSGKGASASTNPADESAAGBNLTEISKKITDSNAFVLAVKEVETIVLSIDELAKKAIG 73
OY      255 OKIDNNGLAALNNONGSLAGAVAIISTLITEKLSKLNLEELKTEIAKAKKCEEFYTK 314
Db      74  OKIDNNGLAALNNONGSLAGAVAIISTLITEKLSKLNLEELKTEIAKAKKCEEFYTK 133
OY      315 LKSGHADLGKODATDDHAKAAILKTHATTDKGAKKEFKDLFESVEGLKKAQVATNSYKE 374
Db      134 LKSGHADLGKODATDDHAKAAILKTHATTDKGAKKEFKDLFESVEGLKKAQVATNSYKE 193
OY      375 L 375
Db      194 L 194

```

Search completed: March 18, 2002, 10:10:56  
 Job time: 979 sec

**THIS PAGE BLANK (USPTO)**



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☒ **OTHER:** page opposite of each other

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

